





```

1 EARLIER FILING DATE: 1997-08-22
2 EARLIER APPLICATION NUMBER: 60/056,631
3 EARLIER FILING DATE: 1997-08-22
4 EARLIER APPLICATION NUMBER: 60/056,845
5 EARLIER FILING DATE: 1997-08-22
6 EARLIER APPLICATION NUMBER: 60/056,892
7 EARLIER FILING DATE: 1997-08-22
8 EARLIER APPLICATION NUMBER: 60/057,761
9 EARLIER FILING DATE: 1997-08-22
10 EARLIER APPLICATION NUMBER: 60/047,595
11 EARLIER FILING DATE: 1997-05-23
12 EARLIER APPLICATION NUMBER: 60/047,599
13 EARLIER FILING DATE: 1997-05-23
14 EARLIER APPLICATION NUMBER: 60/047,588
15 EARLIER FILING DATE: 1997-05-23
16 EARLIER APPLICATION NUMBER: 60/047,594
17 EARLIER FILING DATE: 1997-05-23
18 EARLIER APPLICATION NUMBER: 60/047,589
19 EARLIER FILING DATE: 1997-05-23
20 EARLIER APPLICATION NUMBER: 60/047,593
21 EARLIER FILING DATE: 1997-05-23
22 EARLIER APPLICATION NUMBER: 60/047,614
23 EARLIER FILING DATE: 1997-05-23
24 EARLIER APPLICATION NUMBER: 60/043,578
25 EARLIER FILING DATE: 1997-04-11
26 EARLIER APPLICATION NUMBER: 60/043,576
27 EARLIER FILING DATE: 1997-04-11
28 EARLIER APPLICATION NUMBER: 60/047,501
29 EARLIER FILING DATE: 1997-05-23
30 EARLIER APPLICATION NUMBER: 60/043,670
31 EARLIER FILING DATE: 1997-04-11
32 EARLIER APPLICATION NUMBER: 60/056,632
33 EARLIER FILING DATE: 1997-08-22
34 EARLIER APPLICATION NUMBER: 60/056,664
35 EARLIER FILING DATE: 1997-08-22
36 EARLIER APPLICATION NUMBER: 60/056,876
37 EARLIER FILING DATE: 1997-08-22
38 EARLIER APPLICATION NUMBER: 60/056,881
39 EARLIER FILING DATE: 1997-08-22
40 EARLIER APPLICATION NUMBER: 60/056,909
41 EARLIER FILING DATE: 1997-08-22
42 EARLIER APPLICATION NUMBER: 60/056,875
43 EARLIER FILING DATE: 1997-08-22
44 EARLIER APPLICATION NUMBER: 60/056,862
45 EARLIER FILING DATE: 1997-08-22
46 EARLIER APPLICATION NUMBER: 60/056,887
47 EARLIER FILING DATE: 1997-08-22
48 EARLIER APPLICATION NUMBER: 60/056,908
49 EARLIER FILING DATE: 1997-08-22
50 EARLIER APPLICATION NUMBER: 60/048,964
51 EARLIER FILING DATE: 1997-06-06
52 EARLIER APPLICATION NUMBER: 60/057,650
53 EARLIER FILING DATE: 1997-09-05
54 EARLIER APPLICATION NUMBER: 60/056,884
55 EARLIER FILING DATE: 1997-08-22
56 EARLIER APPLICATION NUMBER: 60/057,669
57 EARLIER FILING DATE: 1997-09-05
58 EARLIER APPLICATION NUMBER: 60/049,610
59 EARLIER FILING DATE: 1997-06-13
60 EARLIER APPLICATION NUMBER: 60/061,060
61 EARLIER FILING DATE: 1997-10-02
62
63 Query Match 4.8%; Score 43; DB 4; Length 1821;
64 Best Local Similarity 100.0%; Pred. No. 4e-08;
65 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0
66
67 856 GCGTAAAAA..... 898

```

[illegible]

SEQUENCE CHARACTERISTICS:  
LENGTH: 740 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-713-000-8

Query Match 4.7%; Score 42; DB 2; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
DB 689 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730

RESULT 5  
US-08-975-316-8  
Sequence 8, Application US/08975316  
Patent No. 5952486

GENERAL INFORMATION:  
APPLICANT: BLOKSBERG, Leonard N., HAYUKKALA, Ilkka  
APPLICANT: and GRIERSON, Alastair W.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR  
THE MODIFICATION OF PLANT LIGNIN CONTENT  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,316  
FILING DATE:  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/713,000  
FILING DATE: September 11, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: SLEATH, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000/1003C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 740 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-975-316-8

Query Match 4.7%; Score 42; DB 2; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
DB 689 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730

RESULT 6  
US-09-211-710-8  
Sequence 8, Application US/09211710A

Patent No. 6204434  
GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
APPLICANT: Hayukkala, Ilkka  
APPLICANT: Grierson, Alastair  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003C3  
CURRENT APPLICATION NUMBER: US/09/211,710A  
CURRENT FILING DATE: 1998-12-14  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 740  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-211-710-8

Query Match 4.7%; Score 42; DB 4; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
DB 689 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730

RESULT 7  
US-09-615-192A-8  
Sequence 8, Application US/09615192A  
Patent No. 6410718

GENERAL INFORMATION:  
APPLICANT: Bloksberg, Leonard N.  
APPLICANT: Hayukkala, Ilkka  
TITLE OF INVENTION: Materials and Methods for the  
FILE REFERENCE: 11000.1003C4U  
CURRENT APPLICATION NUMBER: US/09/615,192A  
CURRENT FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 08/975,316  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: US 08/713,000  
PRIOR FILING DATE: 1996-09-11  
PRIOR APPLICATION NUMBER: US 09/169,789  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 405  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 740  
TYPE: DNA  
ORGANISM: Pinus radiata  
US-09-615-192A-8

Query Match 4.7%; Score 42; DB 4; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
DB 689 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730

RESULT 8  
US-08-975-316-58  
Sequence 58, Application US/08975316  
Patent No. 5952486

GENERAL INFORMATION:  
APPLICANT: BLOKSBERG, Leonard N., HAYUKKALA, Ilkka  
APPLICANT: and GRIERSON, Alastair W.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR  
THE MODIFICATION OF PLANT LIGNIN CONTENT  
NUMBER OF SEQUENCES: 88  
CORRESPONDENCE ADDRESSES:





US-08-655-352-10  
; Sequence 10, Application US/08655352  
; Patent No. 6077991  
; GENERAL INFORMATION:  
; APPLICANT: Bachettira W. Poovaiiah, Zhina Liu,  
; APPLICANT: Shameekumar Patil, Daisuke Takezawa  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &  
; ADDRESSEE: Whinston, LLP  
; STREET: One World Trade Center  
; STREET: 121 S.W. Salmon Street  
; STREET: Suite 1600  
; CITY: Portland  
; STATE: Oregon  
; COUNTRY: United States of America  
; ZIP: 97204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk, 3-1/2 inch  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/655,352  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/323,449  
; FILING DATE: October 14, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Alan. E.  
; REGISTRATION NUMBER: 35,123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (503) 226-7391  
; TELEFAX: (503) 228-9446  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1776 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA to mRNA  
; DESCRIPTION: Tobacco CcAMK cDNA and deduced amino-acid  
; DESCRIPTION: sequence  
; FEATURE:  
; NAME/KEY: protein-coding sequence (not including  
; NAME/KEY: stop codon)  
; LOCATION: nucleotides 20-1570  
US-08-655-352-10  
  
Query Match 4.7%; Score 42; DB 3; Length 1776;  
Best Local Similarity 100.0%; Pred. No. 9.7e-08;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 857 GGTAAAAA  
Db 1719 GGTAAAAA  
RESULT 13  
US-09-258-016-10  
; Sequence 10, Application US/09258016  
; Patent No. 6362395  
; GENERAL INFORMATION:  
; APPLICANT: Bachettira W. Poovaiiah, Zhina Liu,  
; APPLICANT: Shameekumar Patil, Daisuke Takezawa  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Klarquist Sparkman Campbell Leigh &  
ADDRESSEE: Whinston, LLP  
STREET: One World Trade Center  
STREET: 121 S.W. Salmon Street  
STREET: Suite 1600  
CITY: Portland  
STATE: Oregon  
COUNTRY: United States of America  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk, 3-1/2 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/258,016  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Stephens Jr., Donald L.  
REGISTRATION NUMBER: 34,022  
REFERENCE/DOCKET NUMBER: 4630-51994  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 226-7391  
TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1776 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double stranded  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
DESCRIPTION: Tobacco CcAMK cDNA and deduced amino-acid  
DESCRIPTION: sequence  
FEATURE:  
NAME/KEY: protein-coding sequence (not including  
NAME/KEY: stop codon)  
LOCATION: nucleotides 20-1570  
US-09-258-016-10  
  
Query Match 4.7%; Score 42; DB 4; Length 1776;  
Best Local Similarity 100.0%; Pred. No. 9.7e-08;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 857 GGTAAAAA  
Db 1719 GGTAAAAA  
RESULT 14  
US-09-257-8258-10  
; Sequence 10, Application US/092578258  
; Patent No. 6403352  
; GENERAL INFORMATION:  
; APPLICANT: Poovaiiah, Bachettira W.  
; APPLICANT: Patil, Shameekumar  
; APPLICANT: Takezawa, Daisuke  
; TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants  
; FILE REFERENCE: 4630-51993  
; CURRENT APPLICATION NUMBER: US/09/257,8258  
; PRIOR FILING DATE: 1999-02-25  
; PRIOR APPLICATION NUMBER: US 08/655,352  
; PRIOR FILING DATE: 1996-05-23  
; PRIOR APPLICATION NUMBER: US 60/014,743  
; PRIOR FILING DATE: 1996-03-28  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 1776  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
US-09-257-8258-10

Query Match 4.7%; Score 42; DB 4; Length 1776;  
Best Local Similarity 100.0%; Pred. No. 9.7e-08;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
DB 1719 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1760

## RESULT 15

US-08-606-505B-1  
; Sequence 1, Application US/08606505B  
; Patent No. 6114601

## GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro  
APPLICANT: KIYOKAWA, Shigeto  
APPLICANT: SHIMADA, Yukinisa  
APPLICANT: OHBAYASHI, Masaya  
APPLICANT: SHIMADA, Ritsuko  
APPLICANT: OKINAKA, Yasushi  
TITLE OF INVENTION: NOVEL PLANT GENES  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
STREET: 30 Rockefeller Plaza  
CITY: New York

STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112-3801

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/606,505B

FILING DATE: 23-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JPA4963/92

FILING DATE: 02-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Peery, Lawrence S.

REGISTRATION NUMBER: 31865

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-218-2100

TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 1 :

SEQUENCE CHARACTERISTICS:

LENGTH: 1824 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Petunia hybrida

STRAIN: Falcon Blue

FEATURE:

NAME/KEY: CDS

LOCATION: 116 to 1633

IDENTIFICATION METHOD: by experiment

US-08-606-505B-1

Query Match 4.7%; Score 42; DB 3; Length 1824;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 897  
|||||  
DB 1783 GGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1824

Search completed: February 22, 2003, 06:08:15  
Job time : 82.4172 secs



RESULT	1
LOCUS	BE512404
DEFINITION	BE512404 486 bp mRNA linear EST 07-AUG-2000 946070F05.y1 946 - tassell primordium prepared by Schmidt lab Zea
ACCESSION	mays cDNA, mRNA sequence.
VERSION	BE512404
KEYWORDS	BE512404.1 GI:9733652
SOURCE	BST.
ORGANISM	Zea mays. Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 486)
REFERENCE	walbot.v.
AUTHORS	Maize ESTs from various cDNA libraries sequenced at Stanford University
TITLE	Unpublished (1999) Contact: Walbot V
JOURNAL	Department of Biological Sciences Stanford University
COMMENT	855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946070 row: F column: 05. Location/Qualifiers 1. 486 /organism="Zea mays"
FEATURES	
SOURCE	

```

/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassell primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/notes="Organ: Tassels; Vector: HybridZAP; Site:1: EcorI;
Site:2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."

BASE COUNT      100 a      131 c      150 g      105 t
ORIGIN

Query Match      29.4%; Score 264; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 3e-51;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 600 TGGCCACGGGCGCCACGCTCAGCTCTGCTGTACACCCCGACGCGCAACTGCAGGCG 659
    |||
Db 177 TGGCGCCAGGGGCGCCACGCTCAGCTCTGCTGTACACCCCGACGCGCAACTGCAGGCG 236
    |||

QY 660 CAGAGCCCCCTACTAGCTAGTGTAGTGTATGAGTGTAGGAGGAGCCCAACTGCCGCC 719
    |||
Db 237 CAGAGCCCCCTACTAGCTAGTGTAGTGTATGAGTGTAGGAGGAGCCCAACTGCCGCC 296
    |||

QY 720 GCGGCGGACGAGAGTACAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTAT 779
    |||
Db 297 GCGGCGGACGAGAGTACAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTAT 356
    |||

QY 780 GTTTTGTGTGTATCCGCTGCTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
    |||
Db 357 GTTTTGTGTGTATCCGCTGCTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 416
    |||

QY 840 ATCAGCTTTTGTGCGGGTAA 863
    |||
Db 417 ATCAGCTTTTGTGCGGGTAA 440
    |||

RESULT 2
AM923743      297 bp      mRNA      linear      EST 19-JUL-2000
LOCUS
DEFINITION   Dcl_59_A04_g1_A002 Dark Grown 1 (Dcl) Sorghum bicolor cDNA, mRNA
SEQUENCE
ACCESSION   AM923743
VERSION     AM923743.1 GI:8089568
KEYWORDS
SOURCE
ORGANISM     sorghum.
              Sorghum bicolor
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade; Panicoideae; Andropogoneae; Sorghum.
              1 (bases 1 to 297)
              Cordomier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
              ,L.H.
              An EST database from Sorghum: dark-grown seedlings
              Unpublished (2000)
              Contact: Cordomier-Pratt MM
              Laboratory for Genomics and Bioinformatics
              The University of Georgia, Department of Plant Biology
              Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
              Tel: 706 542 1860
              Fax: 706 583 0210
              Email: mpratt@uga.edu
              Sequences have been trimmed to exclude PolyA, vector and regions
              below Phred quality 16. The threshold for highest quality sequence
              is 20.
              Seq primer: PolYTmX
              High quality sequence start: 16
              POLYA-MO.

```

```

FEATURES
source
Location/Qualifiers
1..297
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (Dcl)"
/notes="Organ: 5-day-old dark-grown seedlings; Vector:
lambda zap; Site:1: XhoI; Site:2: EcorI; The library was
made from poly-A RNA in the cloning vector lambda zap II.
Clones to be sequenced were prepared by mass excision."

BASE COUNT      46 a      86 c      97 g      67 t      1 others
ORIGIN

Query Match      9.4%; Score 84; DB 10; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.1e-10;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 597 AGCTGCGCCACGGGCGCCACGCTCAGCTCTGCTGTACACCCCGACGCGCAACTGCAG 656
    |||
Db 43 AGCTGCGCCACGGGCGCCACGCTCAGCTCTGCTGTACACCCCGACGCGCAACTGCAG 102
    |||

QY 657 GGGCAGAGCCCTACTACTAGCT 680
    |||
Db 103 GGGCAGAGCCCTACTACTAGCT 126
    |||

RESULT 3
BM101325      362 bp      mRNA      linear      EST 23-JUL-2002
LOCUS
DEFINITION   Ebp101_S0003_113_R pistil, 1 DPA, no treatment, cv Optic, Ebp101
              Hordeum vulgare cDNA clone Ebp101_S0003_113 5', mRNA sequence.
ACCESSION   BM101325
VERSION     BM101325.2 GI:21944137
KEYWORDS
SOURCE
ORGANISM     Hordeum vulgare.
              Hordeum vulgare
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
              ; Triticeae; Hordeum.
              1 (bases 1 to 362)
              Ramsay,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
              Hensley,L., Machray,G., Marshall,D.F.M. and Waugh,R.
              Development of barley Transcriptome Resources
              Unpublished (2001)
              On Nov 21, 2001 this sequence version replaced gi:17032393.
              Contact: Waugh R, Marshall DF
              Genome Dynamics/Computational Biology
              Scottish Crop Research Institute
              Invergowrie, Dundee, DD2 5DA, Scotland, UK
              Tel: 00 44 1382 562731
              Fax: 00 44 1382 562426
              Email: estescr1.sari.ac.uk
              All sequence has a Phred quality score of 20 or over
              Seq primer: M13 reverse.
              Location/Qualifiers
              1..362
              /organism="Hordeum vulgare"
              /cultivar="Optic"
              /db_xref="taxon:4513"
              /clone="Ebp101_S0003_113"
              /clone_lib="pistil, 1 DPA, no treatment, cv Optic, Ebp101"
              /tissue_type="pistil"
              /dev_stage="1 DPA"
              /lab_host="DH10B"
              /note="Vector: pSPORT1; Site:1: Sal I; Site:2: Not I;
              Non-normalised library, directionally cloned into pSPORT1.
              Derived from pistils dissected from developing grains (24
              hours post anthesis) in glasshouse grown barley plants.
              Developed as part of the barley transcriptome resources of
              BBSRC/SEERAD funded cereal IGF (Investigating Gene
              Function) project."

BASE COUNT      57 a      139 c      129 g      37 t
ORIGIN

```

Query Match	5.6%	Score 50	DB 13	Length 362
Best Local Similarity	100.0%	Pred. No.	0.013	
Matches	50	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0
QY	408	GCACGCCCCCTACGAGGCGGAGCAACGAGGAGTGGGAGACTGACGCGCGGCGCCCC	457	
Db	98	GCACGCCCCCTACGAGGCGGAGCAACGAGGAGTGGGAGACTGACGCGCGCCCC	147	

RESULT	LOCUS	DEFINITION	ACCESSTION
4	BG343299	996 bp	linear
	BG343299	996 bp	linear
	HCMSMG0005F16f	Hordeum vulgare pre-anthesis spike EST library	EST 22-OCT-2001
	HCMSMG0008	(white to yellow anther) Hordeum vulgare cDNA clone	
	HCMSMG0005F16f,	mRNA sequence.	
	bg343300		

ACCESSION	DOJ43299	GI:13155628
VERSION	B6343299.1	
KEYWORDS	EST.	
SOURCE	Hordeum vulgare.	
ORGANISM	Hordeum vulgare	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 996)	Wang, R., Close, T. J., Kleinof, A., Wise, R., Begum, D., Fritch, D., Y. Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D. W., Fenton, R. D., Close, S. J., Gates, R. and Main, D.	Development of a genetically and physically anchored EST resource for barley genomics: Morex pre-anthesis spike cDNA library	Unpublished (2001)	Contact: Wang RA

Clemson University genomics institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: [twingeclemson.edu](mailto:twingeclemson.edu)  
 Total hg bases = 471  
 Seq primer: AATTACCTCCTCAGTAACGG  
 High quality sequence stop: 717.  
 Location/Qualifiers  
 1..996

organism="Hordium vulgare"  
cultivar="Morex"  
/db\_xref="taxon:4513"  
clone="HVSMeq0005516f"  
/clone.lib="Hordium vulgare pre-anthesis spike EST library"  
HYCDNA0008 (white to yellow anther)"  
/tissue\_type="pre-anthesis spike"  
/lab\_host="SOLR"  
/note="vector: lambdaZAP; site\_1: EcoRI; site\_2: XhoI;  
plants were grown in the greenhouse at the University of  
California, Riverside (Fenton, SJ Close, TJ Close). Whole  
spike with awns trimmed were collected at white, green and  
yellow anther stages (Fenton). Total RNA was prepared from  
each pool, equal quantities of all three RNA pools were  
combined, poly(A) RNA was purified from the mixture, one  
primary unamplified cDNA library was made, and 1 million  
pfu were in vivo excised to give plasmidscript SK(-) cDNA  
phagemids. These steps were performed in the TJ Close lab  
(Choi) at the University of California, Riverside.  
Phagemids were plated and picked at the Clemson University  
Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins  
and Wing) Plasmid DNA preparations, DNA sequencing and  
sequence analysis were performed at CUGI (Wing, Yu, Frisch  
Henry, Simmons, Oates, Rambo, Main). The sequence has  
been trimmed to remove vector sequence and contains a  
minimum of 100 bases of phased value 20 or above. For more  
details on library preparation and sequence analysis see  
<http://www.genome.clemson.edu/projects/barley>. To order  
this clone see <http://www.genome.clemson.edu/orders> Also  
see Close TJ, Wing R, Kleinborts A, Wise R (2001)  
genetically and physically anchored EST resources for

BASE COUNT	ORIGIN
183 a	barley genomics. Barley Genetics Newsletter 31:29-30. ( <a href="http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html">http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html</a> )"
310 c	
377 g	
123 t	
3 others	

Query Match	5.6%;	Score 50;	DB 12;	Length 996;
Best Local Similarity	100.0%;	Pred. No. 0.0049;		
Matches	50;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;
07	408	GGCAGCCCCCTACGGGCGGAGCAACGAGGGGTGGGCGAGACGTACCGCGCGCGCC	457	
Db	399	GGCAGCCCCCTACGGGCGGAGCAACGAGGGGTGGGCGAGCTACCGCGCGCC	448	

RESULT 5	
BE420593	
LOCUS	
DEFINITION	
ACCESION	
BE420593	1243 bp mRNA
HMM000.E06	IITEC HMM barley leaf library
HMM000.E06	mRNA sequence.
BE420593	linear EST 24-JUL-2000
	Hordeum vulgare cDNA clone

ACCESSION	B6420593
VERSION	B6420593.1
KEYWORDS	GI:9418436
SOURCE	EST.
ORGANISM	Hordeum vulgare. Hordeum vulgare. Einkorn; Vitis; Vitaceae; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae Triticale; Hordeum. 1 (bases 1 to 1243)
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier
AUTHORS	

**TITLE**  
Expressed Sequence Tags for Species of the Triticaceae  
Published (2000)  
**JOURNAL**  
Contact: Herrmann RG

Botanisches Institut der LMU  
Menzinger Str. 67, D-80638 München GERMANY  
Fax: 49 30 171683  
Email: hermann@botanik.biologie.uni-muenchen.de  
International Triticaceae ESP Cooperative (ITREC)  
<http://wheat.pw.usda.gov/genome>.

FEATURES	SOURCE	LOCATION/QUALIFIERS
1..1243		
/organism="Hordeum vulgare"		
/cultivar="Barke"		
/db_xref="taxon:4513"		
/clone="HMM000.E06"		
/clone.lib="TREC HMM Barley leaf library"		
/tissue.type="leaf"		
/dev_stage="14 day old"		
/note="Vector: pBluescriptSK(-); 850 bp average insert size."		
BASE COUNT	282 a	269 c 364 g 274 t 54 others
ORIGIN		

Query Match	5.3%;	Score 48;	DB 10;	Length 1243;
Best Local Similarity	100.0%;	Pred. No. 0.011;		
Matches 48;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	851	TTGCGCGGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	898	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
DB	615	TTGCGCGGCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	662	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT

RESULT	6
BQ234664/c	
LOCUS	BQ234664
DEFINITION	h48d10.g1 Canis Total Brain CDNAs
ACCESSION	h48d10.5', mRNA sequence.
	BQ234664
	145 bp
	mRNA linear
	EST 03-MAY-2002
	Canis familiaris
	CDNA clone

```

VERSION      BQ234664.1  GI:20430537
KEYWORDS     EST.
SOURCE       dog.
ORGANISM     Canis familiaris
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE    1 (bases 1 to 145)
AUTHORS      O'Shaughnessy,A.L., Palmer,L., McCombie,W.R., Baker,J.P., Bahret,A.,
              Cunniff,D., Dedhia,N.N., de la Bastide,M., Katzenberger,F., King
              L., Kirchoff,K.A., Miller,B., Muller,S., Nascimento,L.U., Preston
              R.R., Shah,R.S., Spiegel,L.A., Zutavern,T., Santos,L. and Hannon
              G.J.
TITLE        Expressed sequence tags from Canis familiaris (dog) (5_2002)
JOURNAL      Unpublished (2002)
COMMENT      Contact: W. Richard McCombie
              Lita Annenberg Hazen Genome Sequencing Center
              Cold Spring Harbor Laboratory
              PO Box 100, Cold Spring Harbor, NY 11724, USA
              Tel: 516 367 8884
              Fax: 516 367 8874
              Email: mcombie@cshl.org
              Plate: hd48 row: d column: 10
              Seq primer: -21M13univrev
              High quality sequence stop: 145.
FEATURES     location/Qualifiers
              1..145
              /organism="Canis familiaris"
              /db_xref="taxon:9615"
              /clone="hd48d10"
              /note="Vector: Lambda Zap II; The library was provided by
              Greg Hannon and Lee Santos (Cold Spring Harbor Laboratory
              ). This library is oligo(dT) primed using stratagene zap
              cdna synthesis kit. It was made from dog whole brain
              cells. Please contact Greg Hannon (hannon@cshl.org) with
              any library related inquiries."
BASE COUNT   14 a 14 c 19 g 98 t
ORIGIN
Query Match 5.0%; Score 45; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 854 CCGGGTAAAAA 898
Db 109 CCGGGTAAAAA 65
RESULT 7
LOCUS      A1452556 178 bp mRNA linear EST 13-APR-1999
DEFINITION tJ22a03.x1 NCI-CGAP Gas4 Homo sapiens CDNA clone IMAGE:2142220 3'
            similar to contains element OFR repetitive element; mRNA
            sequence.
ACCESSION  A1452556
VERSION    A1452556
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 178)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs.fda.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center

```

```

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 764 Std Error: 0.00
Seq primer: -40UP from Gldco
High quality sequence stop: 166.
FEATURES     Location/Qualifiers
              1..178
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:2142220"
              /clone_lib="NCI-CGAP Gas4"
              /tissue_type="poorly differentiated adenocarcinoma with
              signet ring cell features"
              /lab_host="DH10B"
              /note="Organ: stomach; Vector: pCMV-SPORT6; Site: 1; Salt;
              Site: 2; NctI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.69 kb. Life Technologies catalog #:
              11549-011"
BASE COUNT   34 a 40 c 16 g 88 t
ORIGIN
Query Match 5.0%; Score 45; DB 9; Length 178;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 854 CCGGGTAAAAA 898
Db 70 CCGGGTAAAAA 26
RESULT 8
LOCUS      BE429386 180 bp mRNA linear EST 26-JUL-2000
DEFINITION MTD017.E07E990621 TTEC MTD Durum wheat Root Library Triticum
            turgidum subsp. durum CDNA clone MTD017.E07, mRNA sequence.
ACCESSION  BE429386
VERSION    BE429386
KEYWORDS   EST.
SOURCE     durum wheat.
ORGANISM   Triticum turgidum subsp. durum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 180)
AUTHORS    Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
            S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
            Herrmann,R.G., Holton,T., Jacquemlin,J.M., Jia,J., Joudrier,P.,
            Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
            Recchioni,N., Qualset,C., Schuch,M., Selvaraj,G., Shariflou,M.,
            Sorrells,M., Wardurton,M. and Wenzel,G.
            International Triticeae EST Cooperative (TTEC): Production of
            Expressed Sequence Tags for Species of the Triticeae
            Unpublished (2000)
            Contact: Joudrier P
            INRA, Unite de Biochimie et Biologie Molculaire des Cereales
            2, place VIALA, 34060 Montpellier cedex 01 FRANCE
            Tel: 33 4 99 61 23 84
            Fax: 33 4 99 61 23 48
            Email: joudrier@enscm.inra.fr
            International Triticeae EST Cooperative (TTEC)
            http://wheat.pw.usda.gov/genome.
FEATURES     Location/Qualifiers
              1..180
              /organism="Triticum turgidum subsp. durum"
              /cultivar="Silihana"
              /db_xref="taxon:4567"
              /clone="MTD017.E07"
              /clone_lib="TTEC MTD Durum Wheat Root Library"
              /tissue_type="root"
              /dev_stage="3-day-old seedling, water-stressed"
              /note="Vector: pSPORT1; T7 primers used. See pSPORT1
              polylinker site. 0.3-2.0 kbp average insert size."

```



BASE COUNT 61 a 20 c 63 g 14 t 22 others  
ORIGIN

Query Match 5.0%; Score 45; DB 10; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.34;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 854 CCGGGTAAAAA 898  
Db 1 CCGGGTAAAAA 45

RESULT 9  
Bg122481 192 bp mRNA linear EST 30-JAN-2001  
LOCUS 602353281F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4451253 5',  
DEFINITION mRNA sequence.  
ACCESSION Bg122481  
VERSION Bg122481.1 GI:12615990  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 192)  
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM10238 row: b column: 22  
High quality sequence stop: 145.  
Location/Qualifiers  
1. .192  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4451253"  
/clone\_lib="NIH\_MGC\_90"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; Oligo-dt primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."

BASE COUNT 143 a 8 c 22 g 19 t  
ORIGIN

Query Match 5.0%; Score 45; DB 12; Length 192;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 854 CCGGGTAAAAA 988  
Db 52 CCGGGTAAAAA 96

RESULT 10  
AM169684 229 bp mRNA linear EST 12-NOV-1999  
LOCUS x130107.x1 NCI\_CGAP\_Ut2 Homo sapiens cDNA clone IMAGE:2658781 3',  
DEFINITION mRNA sequence.  
ACCESSION AM169684  
VERSION AM169684.1 GI:6401209  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 229)  
REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

FEATURES  
source  
1. .229  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2658781"  
/clone\_lib="NCI\_CGAP\_Ut2"  
/tissue\_type="moderately-differentiated endometrial  
adenocarcinoma, 3 pooled tumors"  
/lab\_host="DH10B"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.85 kb. Life Technologies catalog #: 11599-012"

BASE COUNT 55 a 46 c 40 g 88 t  
ORIGIN

Query Match 5.0%; Score 45; DB 10; Length 229;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 854 CCGGGTAAAAA 898  
Db 59 CCGGGTAAAAA 15

RESULT 11  
BI002437 300 bp mRNA linear EST 13-JUN-2001  
LOCUS MR3-HN0150-220101-007-h08 HN0150 Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
ACCESSION BI002437  
VERSION BI002437.1 GI:14406511  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 300)  
REFERENCE Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
JOURNAL  
MEDLINE  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Pudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR3&t2=MR3-HN0150-220101-007-h08&t3=2001-01-22&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 276.

## FEATURES

source

location/Qualifiers

1..300

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_id="HN0150"

/dev\_stage="Adult"

/note="Organ: head,normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESMPS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

50 a

60 c

41 g

149 t

## ORIGIN

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.21; Length 300;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGGTAAAAA  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
COMMENT  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA library preparation: Life Technologies, Inc.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Insert length: 929 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 329  
POLY-A-NO.

## FEATURES

source

location/Qualifiers

1..309

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_id="IMAGE:2070196"

/clone\_id="Soares\_NHMPU\_S1"

/tissue\_type="Pooled human melanocyte, fetal heart, and pregnant uterus"

/lab\_host="DH10B"

/note="Organ: mixed (see below); Vector: pTZ19-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NDHM, pregnant uterus NBHPU, and fetal heart NBH19M) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

## BASE COUNT

73 a

58 c

78 g

100 t

## ORIGIN

## Query Match

5.0%; Score 45; DB 9; Length 309;

Best Local Similarity 100.0%; Pred. No. 0.2; Length 309;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGGTAAAAA  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
COMMENT  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA library preparation: Life Technologies, Inc.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Insert length: 929 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 329  
POLY-A-NO.

## FEATURES

source

location/Qualifiers

1..340

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_id="IMAGE:2187329"

/clone\_id="NCI CGAP Gas4"

/tissue\_type="Poorly differentiated adenocarcinoma with signet ring cell features"

/lab\_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.69 kb. Life Technologies catalog #:

## BASE COUNT

63 a

74 c

95 g

108 t

## ORIGIN

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.19; Length 340;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGGTAAAAA  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
COMMENT  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA library preparation: Life Technologies, Inc.  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Insert length: 929 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 329  
POLY-A-NO.

RESULT 14  
 AM151031 355 bp mRNA linear EST 03-NOV-1999  
 LOCUS xq43q09.x1 NCI\_CGAP\_Utl Homo sapiens cDNA clone IMAGE:2630368 3',  
 DEFINITION mRNA sequence.  
 ACCESSION AM151031  
 VERSION AM151031.1 GI:6198929  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 355)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -400P from Glibco  
 High quality sequence stop: 345.  
 Location/Qualifiers  
 1..355  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2630368"  
 /clone\_lib="NCI\_CGAP\_Utl"  
 /tissue\_type="well-differentiated endometrial  
 adenocarcinoma, 7 pooled tumors"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: PCMV-SPORT6; Site:1: SalI;  
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

BASE COUNT 88 a 64 c 100 g 102 t 1 others  
 ORIGIN

Query Match 5.0%; Score 45; DB 10; Length 355;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 58 CCGGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14

RESULT 15  
 AM264516 457 bp mRNA linear EST 28-DEC-1999  
 LOCUS xrl0e06.x1 NCI\_CGAP\_Brn53 Homo sapiens cDNA clone IMAGE:2759170 3',  
 DEFINITION similar to SW:PF12.PIG P51525 PROPHENIN-2 PRECURSOR; contains  
 element MSRI repetitive element; mRNA sequence.  
 ACCESSION AM264516  
 VERSION AM264516.1 GI:6641332  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 457)  
 NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute / National Institute of Neurological  
 Disorders and Stroke, Brain Tumor Genome Anatomy Project

JOURNAL  
 COMMENT  
 (CGAP/BRGAP), Tumor Gene Index  
 Unpublished (1998)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Chris Moskalkuk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
 Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The  
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

FEATURES  
 source  
 1..457  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2759170"  
 /clone\_lib="NCI\_CGAP\_Brn53"  
 /tissue\_type="three pooled meningiomas"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: PCMV-SPORT6; Site:1: SalI;  
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies."

BASE COUNT 98 a 117 c 107 g 135 t  
 ORIGIN

Query Match 5.0%; Score 45; DB 10; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 98 CCGGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54

Search completed: February 22, 2003, 06:06:21  
 job time : 1474.72 secs



score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

tribution.

Qu

(without alignments)  
10485.621 Million cell updates/sec

Sequence: 1 ctgcgacgactcgacgtc.....aaaaaaaaaaaaaaaa 898

Gapop 60.0 , Gapext 60.0

Total number of hits satisfying chosen parameters: 128065

Post-processing: Listing first 45 summaries

```
1:  gb_ba:*
```

40:	em_hyo_hum:*
39:	em_hyo_mus:*
41:	em_hyo_other:*
38:	em_hyo_vrt:*
37:	em_hyo_vrt:*
36:	em_hyo_mam:*
35:	em_hyo_rtd:*
34:	em_hyo_pln:*
33:	em_hyo_mus:*
32:	em_hyo_other:*
31:	em_hyo_inv:*
30:	em_hyo_hum:*
29:	em_v1:*
28:	em_uni:*
27:	em_strs:*
26:	em_ro:*
25:	em_pl:*
24:	em_ph:*
23:	em_pat:*
22:	em_ov:*
21:	em_or:*
20:	em_com:*
19:	em_uni:*
18:	em_in:*
17:	em_hum:*
16:	em_fun:*
15:	em_ba:*
14:	em_v1:*
13:	em_uni:*
12:	em_strs:*
11:	em_sy:*
10:	em_ro:*
9:	em_pl:*
8:	em_pi:*
7:	em_ph:*
6:	em_pat:*
5:	em_ov:*
4:	em_com:*
3:	em_in:*
2:	em_hyo:*

Pred. No. is the number of results predicted by chance to have a

[illegible]

## ALIGNMENTS

RESULT	1
BC019562	
LOCUS	2445 bp
DEFINITION	Homo sapiens, clone IMAGE:3451144, mRNA, partial cds.
ACCESSION	BC019562
VERSION	BC019562.1 GI:18042964
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.

REFERENCE 1 (bases 1 to 2445)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian

REMARK  
COMMENT

Gene Collection (MGC), Cancer Genomics office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amgdbcm.tmc.edu](mailto:amgdbcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAC Plate: 19 Row: c Column: 17.

## FEATURES

## source

Location/Qualifiers

1..2445

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3451144"

/tissue\_type="Placenta, choriocarcinoma"

/clone\_id="NHL\_MGC\_10"

/lab\_host="DH10B"

/note="vector: pCMV-SPORT6"

<1..1124

/codon\_start=3

/product="Unknown (protein for IMAGE:3451144)"

/protein\_id="AAH19562.1"

/db\_xref="GI:18042965"

/translation="IKLIDAGVSLIVLVVAGSGGLVFLNCFNPLPPEPEMD

DYVKIKRSMGEPHRTTKQRYKQVTVGRRLSSSSMRSAKEVALQEBHKLCT

VDLEVKCPDAVAVPFSVSPFILLSTVTCVQLRITMGAMNNILFLVSGD

QKTVGLTSLFVGLCLTLPVIGYIMDWRLKEDESEPEERDANQGEKKRR

DRIOIKITNARAFATNLLVGFVTCIPNLPQILSPILHTIVGFHSVAGGLY

AAVPSPTGSLTGLISLALFALQCPFLAMGFLQGDPLVNVGLLLSLGFC

LPLVLCRRRLERLOOQROEDDKFLKINGSSNQEAFF"

## CDS

## BASE COUNT

498 a 765 c 702 g 480 t

## ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 45; DB 9; Length 2445;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 CCGGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898

Db 2319 CCGGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2363

RESULT 2  
 LOCUS AY070835 1981 bp mRNA linear INV 20-DEC-2001  
 DEFINITION Drosophila melanogaster GH06691 full length cDNA.  
 ACCESSION AY070835  
 VERSION AY070835.1 GI:17944780  
 KEYWORDS FLI.CDNA.  
 SOURCE Drosophila melanogaster.  
 ORGANISM Drosophila melanogaster.  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1981)

REFERENCE  
AUTHORS

Stephenson, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
 Champagne, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,  
 George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,  
 Miranda, A., Mungall, C. J., Nuno, J., Paclob, J., Paragas, V., Park, S.,  
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.  
 and Celisner, S.

TITLE  
JOURNAL

## COMMENT

Direct Submission  
 Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,  
 Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, USA  
 Sequence submitted by:  
 Lawrence Berkeley National Laboratory  
 Berkeley Drosophila Genome Project  
 This clone was sequenced as part of a high-throughput process to  
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
 Science 2000). The sequence has been subjected to integrity checks  
 for sequence accuracy, presence of a polyA tail and contiguity  
 within 100 kb in the genome. Thus we believe the sequence to  
 reflect accurately this particular cDNA clone. However, there are  
 artifacts associated with the generation of cDNA clones that may  
 have not been detected in our initial analyses such as internal  
 priming, priming from contaminating genomic DNA, retained introns  
 due to reverse transcription of unsliced precursor RNAs, and  
 reverse transcriptase errors that result in single base changes.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our web site  
 (<http://fruitfly.berkeley.edu>) or send email to  
[cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

## FEATURES

## source

Location/Qualifiers

1..1981

/organism="Drosophila melanogaster"

/strain="y; cn bw sp"

/db\_xref="taxon:7227"

/map="67E5-67E5"

1..1981

/gene="CG11811"

/note="alignment with genomic scaffold AE003547"

/db\_xref="FLYBASE:FBgn0036099"

506..1207

/gene="CG11811"

/note="Longest ORF"

/codon\_start=1

/product="GH06691P"

/protein\_id="AAL8457.1"

/db\_xref="GI:17944781"

/db\_xref="FLYBASE:FBgn0036099"

/translation="MISPLFVNALSSSSSSAASLTSKMTAPGRPLVLCGPS

GSCKSLTKRFLFAEPSPGFSISHTKREBEHGVYVEREMAAIAGGEFT

ETAEPTGVLVGTSGRAVRETOAGRCVYIIDIEKGVQIKRDPVLIENPPIKE

LEKRLKRGSTERSLSKRLNAQVEIDYGLTGNNHKLINNDIVDAVEERFVVO

EKEEOQKGVSVNLN"

## BASE COUNT

617 a 438 c 464 g 462 t

## ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 44; DB 3; Length 1981;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 855 CCGGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898

Db 1909 CCGGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1952

RESULT 3  
 LOCUS AX187087 293 bp DNA linear PAT 06-AUG-2001  
 DEFINITION Sequence 2782 from Patent WO0142467.  
 ACCESSION AX187087  
 VERSION AX187087.1 GI:15138532  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.

## REFERENCE

1 (bases 1 to 293)  
 Schleigel, R., Deeds, D., Berger, A. and Zhao, X.  
 Genes, compositions, kits, and methods for identification,  
 assessment, prevention, and therapy of cervical cancer

```

JOURNAL Patent: WO 0142467-A 2782 14-JUN-2001;
          Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
  source 1..293
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
BASE COUNT 54 a 80 c 11 g 148 t
ORIGIN

Query Match 4.8%; Score 43; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 2.4e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GGGTAAAAA... 898
Db 96 GGGTAAAAA... 54

RESULT 4
AF043538 1032 bp mRNA linear PLN 18-AUG-1998
LOCUS Arabidopsis thaliana 20S proteasome beta subunit PBG1 (PBG1) mRNA,
DEFINITION complete cds.
ACCESSION AF043538
VERSION AF043538.1 GI:3421122
KEYWORDS Arabidopsis thaliana.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1032)
AUTHORS Fu,H., Doelling,J.H., Arendt,C.S., Hochstrasser,M. and
Vierstra,R.D.
Molecular organization of the 20S proteasome gene family from
Arabidopsis thaliana
Genetics 149 (2), 677-692 (1998)
JOURNAL MEDLINE 98278790
PUBMED 9611183
REFERENCE 2 (bases 1 to 1032)
AUTHORS Fu,H.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1998) Horticulture, U. of Wisconsin, 1575 Linden
Dr., Madison, WI 53706, USA
FEATURES Location/Qualifiers
  source 1..1032
           /organism="Arabidopsis thaliana"
           /strain="Columbia"
           /db_xref="taxon:3702"
           /clone="147M21T7"
           1..1032
           /gene="PBG1"
           43..783
           /gene="PBG1"
           /EC_number="3.4.99.46"
           /codon_start=1
           /product="20S proteasome beta subunit PBG1"
           /protein_id="AAC32074.1"
           /db_xref="GI:3421123"
           /translation="MTFSPVINDGSMKLAEESSORTIVPYVGTGSAVIAIKKDGVL
MASDMSGSGSTLRKNTKERVKAIGKSLGASGELSDPELIRYDELTLNNMDD
GNSLGRKEHNITLVRYNRRNKRPLNNTIYVGGVANKGSIYAGMSMGIVSFEDDHV
ATGFNSHLARPLTRDEWADDSFEDGVALKCKMKRVLLYRDRSAINKLQIAKITEBSV
TVSPSPSLKTYEFSFNPYTAGAGSW"
BASE COUNT 343 a 178 c 227 g 284 t
ORIGIN

Query Match 4.8%; Score 43; DB 8; Length 1032;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GGGTAAAAA... 898
Db 999 GGGTAAAAA... 1041

RESULT 6
BC004060 1303 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:8174 IMAGE:3590170, mRNA, complete cds.
DEFINITION BC004060
ACCESSION BC004060
VERSION BC004060.1 GI:13278527
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1303)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

```

Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [md@paxil.stanford.edu](mailto:md@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnln.gov>  
Series: IRAC Plate: 11 Row: c Column: 5.

## FEATURES

source

1. 1303

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="CECH 11"

/clone="MGC:8174 IMAGE:3590170"

/tissue\_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone\_id="NCI CGAP\_L429"

/lab\_host="DH10B"

/note="vector: pCMV-SPORT6"

## CDS

162..839

/codon\_start=-1

/product="Unknown (protein for MGC:8174)"

/protein\_id="AAH04060.1"

/db\_xref="GI:13278528"

/translation="MGPRLEGIIPALPLPLT/MTLITGZPLMQLSMDCGDLTDGLV  
VLARSGARCTDMRWLRNVRVAPLPLREELVFPACMLPMIAPCTGCGAPVPCPLPGV  
AHFHITTEQLRFROSSVSGSIFVSAAOFSTYVFGATYRFLFIRGHIGYLCSTRC  
NYMGFPAVCALEHPQWPLLAGYALGVCLPLLLQPLTPDKLYSLPLCMILERTGA  
SETLLCS"

## BASE COUNT

280 a 365 c 324 g 334 t

## ORIGIN

Query Match 4.8% Score 43; DB 10; Length 1303;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
Db 1248 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1290

RESULT 7  
LOCUS HSN802527 1445 bp mRNA linear PRI 23-MAR-2000  
DEFINITION Homo sapiens mRNA; cDNA DKFZp586C2017 (from clone DKFZp586C2017).  
ACCESSION AL162008  
VERSION AL162008.1 GI:7328034  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1445)  
Koehler, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.  
Direct Submission  
Submitted (15-MAR-2000) MIPS, Am Kiopterapitz 18a, D-82152  
Martinsried, GERMANY  
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
Sequenced by BMFZ (Biomedical Research Center at the Charite,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
This clone (DKFZp586C2017) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: [clone@rzpd.de](mailto:clone@rzpd.de) Further  
information about the clone and the sequencing project is available  
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

## FEATURES

source

1. 1445

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKFZp586C2017"

/tissue\_type="uterus"  
/clone\_id="586 (synonym: hurel). Vector pSPORT1; host  
DH10B; sites NotI + SalI/MluI"  
/dev\_stage="adult"  
1342..1347  
POLYA\_signal  
POLYA\_site 1361  
BASE COUNT 523 a 206 c 232 g 484 t  
ORIGIN

Query Match 4.8% Score 43; DB 9; Length 1445;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
Db 1358 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1400

## RESULT 8

LOCUS BC005424

DEFINITION Mus musculus, tumor susceptibility gene 101, clone MGC:5785  
IMAGE:3490301, mRNA, complete cds.

ACCESSION BC005424

VERSION BC005424.1 GI:13529355

KEYWORDS MGC.  
house mouse.ORGANISM Mus musculus.  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.REFERENCE 1 (bases 1 to 1463)  
Straussberg, R.

## AUTHORS

Direct Submission

## JOURNAL

Submitted (27-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaphs-r@mail.nih.gov](mailto:cgaphs-r@mail.nih.gov)

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [md@paxil.stanford.edu](mailto:md@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnln.gov>  
Series: IRAC Plate: 7 Row: j Column: 24  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11230779.  
Location/Qualifiers  
1. 1463

## FEATURES

source

/organism="Mus musculus"

/db\_xref="taxon:10090"

/map="C57BL/6J"

/clone="MGC:5785 IMAGE:3490301"

/tissue\_type="Mammary tumor. WAP-TGF alpha model. 7 months  
old, gross tissue."

/clone\_id="NCI CGAP\_Mam5"

/lab\_host="DH10B"

/note="vector: pCMV-SPORT6"

85..1260

/codon\_start=-1

/product="tumor susceptibility gene 101"  
/protein\_id="AAH05424.1"  
/db\_xref="GI:13529356"  
/db\_xref="LOCUSID:22088"  
/translation="MAVSEQLKMKSKYKRLVYKQTVNVIAMKDKLKVLSYVF"



NDGSSREIVNLTGTFIPVRYGNITVNICPILMLDTYPYNPICFVCKPTSSMTIKTKGH  
VDANGKTYLPIYLDHMKHPRSELLELOIMYIFGEEPRVSRVSAVSPYATAGBP  
NTSYMGMBGSIAYRSGYPPRPSGYPGCPYPPAGPYATTSSQYPPQPYTVYGBSR  
DGTISEDTJASLISAVSKLPMKREEDGAAQALNALKTEEDLKGGOKLEBWT  
RLOEVAEVDKRNLELKKEDELSSALKEKMOSENNDIVELLPTAPLKOILNLTA  
EENAIEDTIFYLEALRNGVIDLVFLKRVLLSRKQFOLRALMOKARKTAGSILDY\*

BASE COUNT 451 a 342 c 313 g 357 t

ORIGIN

Query Match 4.8%; Score 43; DB 10; Length 1463;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 856 GGCTAAAAA 898  
Db 1407 GGCTAAAAA 1449

RESULT 9  
LOCUS AR202048 1835 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 1 from patent US 6361948.  
ACCESSION AR202048  
VERSION AR202048.1 GI:20256587  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1835)  
AUTHORS Tricoll, J.V. and Roudine, J.L.  
TITLE Prognostic compositions for prostate cancer and methods of use thereof  
JOURNAL Patent: US 6361948-A 1 26-MAR-2002;  
FEATURES Location/Qualifiers  
source 1..1835  
BASE COUNT 384 a 607 c 514 g 330 t  
ORIGIN

Query Match 4.8%; Score 43; DB 6; Length 1835;  
Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 856 GGCTAAAAA 898  
Db 1749 GGCTAAAAA 1791

RESULT 10  
LOCUS HS058996 1853 bp mRNA linear PRI 10-OCT-2001  
DEFINITION Homo sapiens testis calpastatin mRNA, complete cds.  
ACCESSION U58996  
VERSION U58996.2 GI:5042381  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1853)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
TITLE Li, S., Liang, Z.G., Wang, G.Y., Yavetz, B., Kim, E.D. and Goldberg, E.  
JOURNAL Molecular cloning and characterization of functional domains of a  
MEDLINE human testis-specific isoform of calpastatin  
20318416  
PUBMED 10859257  
REFERENCE 2 (bases 1 to 1853)  
AUTHORS O'Hern, P.A., Liang, Z.G., Wang, G.Y., Yavetz, B., Kim, E. and  
TITLE A novel testis-specific isoform of calpastatin detected with serum  
JOURNAL from an infertile patient  
REFERENCE Unpublished  
3 (bases 1 to 1853)

AUTHORS O'Hern, P.A., Liang, Z.G., Wang, G.Y., Yavetz, B., Kim, E. and  
TITLE Direct Submission  
JOURNAL Submitted (22-MAY-1996) Department of Biochemistry, Molecular  
Biology and Cell Biology, Northwestern University, 2153 Sheridan  
Road, Evanston, IL 60208, USA  
REFERENCE 4 (bases 1 to 1853)  
AUTHORS Goldberg, E.  
TITLE Direct Submission  
JOURNAL Submitted (15-APR-1998) Department of Biochemistry, Molecular  
Biology and Cell Biology, Northwestern University, 2153 Sheridan  
Road, Evanston, IL 60208, USA  
REFERENCE 5 (bases 1 to 1853)  
AUTHORS Goldberg, E.  
TITLE Direct Submission  
JOURNAL Submitted (21-APR-1998) Department of Biochemistry, Molecular  
Biology and Cell Biology, Northwestern University, 2153 Sheridan  
Road, Evanston, IL 60208, USA  
REMARK Nucleotide and amino acid sequences updated by submitter  
6 (bases 1 to 1853)  
AUTHORS Goldberg, E.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUN-1999) Department of Biochemistry, Molecular  
Biology and Cell Biology, Northwestern University, 2153 Sheridan  
Road, Evanston, IL 60208, USA  
REMARK Sequence update by submitter  
On Jun 11, 1999 this sequence version replaced gi:3068549.  
FEATURES Location/Qualifiers  
source 1..1853  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="testis"  
67..1359  
/function="inhibitor of calpain (calcium-dependent neutral  
protease)"  
/note="c-Ca<sup>2+</sup>; testis specific isoform of calpastatin"  
/protein\_id="AAC14353.1"  
/db\_xref="GI:3068550"  
/translational="MGQFLSTFLESGPPTVWHDKCDGREGAREAVRIJFODAKK  
EERLEKSGEDDETIPSEVRLKPAIDDKKPLPEPERKPRRSSELIDELSDPDS  
ECKEKPEKPEKTEESKAAAPAVSRAVRSKSGIOSNPREPTILGTYDDAVEL  
ADSLGKREAPDEGKPYMDKVKAKAKEDREKRETTTPDTRLEVKNDKQPL  
PKSEKQLPMSDDYFLDALISDFSGPQNASLTKEDAKLAALISEVVSQTPASTQA  
GAPPRDTSQDLDDALDKLSDELQRODPDENKPMEDKVKERAKAHRDKLGERD  
TIPPEYRHLLDDNGODKPYKPTKSEDSKPADQDDIDALISGLDSCPSYETSON  
TAKDKCKKAASSSKAPKNGKAKDSAKTETSKPKD"  
67..190  
misc\_feature  
/note="encodes amino acids unique to testis isoform"  
BASE COUNT 674 a 370 c 412 g 397 t  
ORIGIN

Query Match 4.8%; Score 43; DB 9; Length 1853;  
Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Yy 856 GGCTAAAAA 898  
Db 1771 GGCTAAAAA 1813

RESULT 11  
LOCUS BC003999 1913 bp mRNA linear ROD 07-AUG-2002  
DEFINITION Mus musculus, clone MGC:7540 IMAGE:3492402, mRNA, complete cds.  
ACCESSION BC003999  
VERSION BC003999.2 GI:16306793  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 1913)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://imgc.nci.nih.gov>  
 On Oct 22, 2001 this sequence version replaced gi:13278363.  
 CONTACT: MGC help desk  
 E-mail: gcgaps-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLND)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 CONTACT: (Dickson, Mark) mcdickax@stanford.edu  
 Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLND at: <http://image.llnl.gov>  
 Series: IRAC Plate: 8 Row: h Column: 7.

FEATURES

Location/Qualifiers

1..1913  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="C2E6H 11"  
 /clone="MGC:7540 IMAGE:3492402"  
 /tissue\_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."  
 /clone\_id="MCL.CGAP.Lu29"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 129..1205  
 /codon\_start=1  
 /product="unknown (protein for MGC:7540)"  
 /protein\_id="AAH03999.1"  
 /db\_xref="GI:13278364"

CDS

/translation="MAP0LSPFCILLIYIGTVIAGDFYKIIGVPSASINDIKK  
 YKLAQLHPRNPDPPQAEKFDIGAAVEYLSERKQIDTYGEGGLDQSH  
 GDIFSHFDFGFMGTGPRDODRINPRSDIIVLEVTLEVANGFEVNRKPA  
 ROAPGRKCKRQEMRTQLGPRFQMTQEVVDCDPCPVKLVNERTLEVEIEPGVD  
 GMEYPTGEGEPHVDGDELRIRIKVYKHFERRGDLITNTVSLVEALVGFEND  
 ITHLDGKVIHSRDKITRPGAKIMKKGELPFDNNINIKGSLITTFVDPEPKQITEE  
 AKGIGKOLKQGPVKVYNGLOGY"

BASE COUNT 576 a 401 c 508 g 428 t  
 ORIGIN

Query Match 4.88; Score 43; DB 10; Length 1913;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
 |||  
 Db 1834 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1876

RESULT 12  
 SSA293335 2668 bp mRNA linear VRT 15-JUL-2002  
 LOCUS SSA293335  
 DEFINITION Salmo salar mRNA for putative interleukin-1 receptor type I (11-111 gene)  
 ACCESSION AJ293335  
 VERSION AJ293335.1 GI:19912978  
 KEYWORDS IL-1RL gene; interleukin-1 receptor-like protein.  
 SOURCE Atlantic salmon.  
 ORGANISM Salmo salar  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;

REFERENCE 1  
 AUTHORS Cunnigham, C., Stansberg, C., Olsen, L., Zou, J., Secombes, C.J. and  
 TITLE Cloning of a Salmo salar interleukin-1 receptor-like cDNA  
 JOURNAL Dev. Comp. Immunol. 26 (5), 415-431 (2002)  
 MEDLINE 21904083  
 REFERENCE 2 (bases 1 to 2668)  
 AUTHORS Subramaniam, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-SEP-2000) Subramaniam S., Molecular Immunology, Sars International Centre for Marine Molecular Biology, Bergen High Technology Centre, Thormohlensgt. 55, Bergen, N-5008, NORWAY

FEATURES

Location/Qualifiers

1..2668  
 /organism="Salmo salar"  
 /db\_xref="taxon:8030"  
 /clone="pSS211"  
 /tissue\_type="head kidney"  
 /dev\_stage="young adult"  
 /lab\_host="E. coli"  
 /country="Norway:Bergen"  
 1..2668  
 /gene="IL-1RL"  
 1..92  
 /gene="IL-1RL"  
 93..1808  
 /gene="IL-1RL"  
 /function="Immune response"  
 /codon\_start=1  
 /product="interleukin-1 receptor-like protein"  
 /protein\_id="CAC83729.1"  
 /db\_xref="GI:19912979"

CDS

/translation="MLGEFYIRFLSMITLLHVSQGEHOGAIEVHHSAGHPOLKCS  
 DYDAEDPOTWTSRGNGNOTLNTSGEVREGVIMFLPHSHNGHYNCESKYPNSM  
 EMKRVLSVDPCPPVPAENRTYSQGSSEVFPEDVDYGLDPTAHIRFKCSYVDK  
 GETTIEDGPKRLVNAETSHAGVITCLVERSLDGGNTATHTOLKVNKEVPLK  
 QVYPRKRETVTVEGSRVELVCSAFVGEGESSEPMFWVDGRHDHIEQLNVTY  
 TILPRSSGVGRSTLSIYEVREFLWVITSCIALNALQDVGFLMDPANSGFYTCL  
 CIGAFSMVILGVAMCCLFEVDVLVLRRLRPLVAKRPAQGLTDAVSYVHSDGS  
 RAEFALQVPLEVLEERYGYTLFVSGRDLPEEAIHDVTSERMSRRLITLSAOSV  
 SLHPKRRDPDOLPLHQSPQDSYPOOIGLVDALONGLRVYLVEIDGKYDYSLSK  
 SLVYTRKQCALMWRNRSSKSSSTAHPRNGHWKCLRHMPEKPKGLQPLTFCINC  
 STERY"

sig\_peptide 93..155  
 /gene="IL-1RL"  
 mat\_peptide 156..1805  
 /gene="IL-1RL"

3'UTR 1809..2668  
 /product="interleukin-1 receptor-like protein"

polyA\_signal 2594..2600  
 /gene="IL-1RL"  
 polyA\_site 2619  
 /gene="IL-1RL"

BASE COUNT 734 a 577 c 645 g 712 t  
 ORIGIN

Query Match 4.88; Score 43; DB 5; Length 2668;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
 |||  
 Db 2615 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2657

RESULT 13  
 AF069757 3280 bp DNA linear INV 28-MAY-1999  
 LOCUS AF069757  
 DEFINITION dictyostellium discoideum spermidine synthase gene, complete cds.  
 ACCESSION AF069757  
 VERSION AF069757.1 GI:4903272



reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our web site  
(<http://fruitfly.berkeley.edu>) or send email to  
[cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

## FEATURES

## source

## gene

## CDS

1. 5012  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/map="100D2-100D2"  
1. 5012  
/gene="CG1815"  
/note="alignment with genomic scaffold AE003779"  
/db\_xref="FlyBASE:FBgn0039863"  
160. 2883  
/gene="CG1815"  
/note="longest ORF"  
/codon\_start=1  
/product="SD07276p"  
/protein\_id="AAL3986.1"  
/db\_xref="GI:17862984"  
/db\_xref="FlyBASE:FBgn0039863"  
/translation="MAFEDGEFLANLMAEDRNIVPAYMATRCRSRKTLSNSMGN  
DDGPRRSNLRSENEEFKHNAPLDRIIHDDRESEALNTSGESSLFSADDTKT  
TTAKTPEEAEARLALIVEPPTPGMDPCWCKGCGKLMPCSKLRSPSYCYRPAT  
TKPDSSMKGCEQVTAAPKRLRNGVSDLLSOLLSPALDPMKHYRGAKHRSLELV  
FPLTKKFTPNPVFSFSLAOCIRNGATQSTDEFLSEVKIOTHNALLIDAGDAKYEQAS  
KAVVVCROFANEIDTCPECYLNANSDDWFKVCRPHLLMLNKLKGFPPWPKAMG  
SSNSTLVNWFEGKHDAFVPAKDFLYSAQPNPTQTSRSARDLAECIREVELIHID  
IKRRIGAFNAPYRTPYDPLEEQOOLEQMPGVAIDRELEPANKTPIQLIRKTD  
DKLSIVKTKATSEKNSDQSPPTKLSKVYVSVTGSCSDSHSNVSNYEVISRS  
GESLTSRCVKLKRSLAKIIVASVETSEAVAPKRKSHLSDSFTSESEHRRKSK  
HARKHNDNDNOIEEAETGQEPKSPITSKONENLRDEENVITENANDTSASPVSA  
SVYSVYELVRRRGVITTKIPRECOQTAEPTAAVPIPIPTAPPKONTPRGNPAAPK  
QSDVERQOELIKKVIPTETIKTEVASEPDDGITEASPAHQPTDOYPLQOETITQ  
PESOMPAAPQPKPVNDAPFEEVRIKEELISEDEMETQSIVSRKLSVPPMLPM  
PPPPPLPRESPATADSVAFVGDITIQVYHRHPLQTRPFCOPHLRLRHOSQLAR  
ILLASHRLRHRPRESLRTHTFNEEKTIVSQPPHRLHQPOTRLCSDPTW"

BASE COUNT 1485 a 1346 c 1135 g 1046 t  
ORIGIN

Query Match 4.8%; Score 43; DB 3; Length 5012;  
Best Local Similarity 100.0%; Pred. No. 2.1e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898  
|||||  
db 4961 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5003

Search completed: February 22, 2003, 05:26:00  
Job time : 2547.4 secs

GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 04:08:26 ; Search time 88.6106 Seconds  
(without alignments)  
5691.782 Million cell updates/sec

Title: US-09-832-320-1

Perfect score: 898  
Sequence: 1 ctgcgacgactgcgacgctc.....aaaaaaaaaaaaaaaaaaaaa 898

Scoring table: OLIGO\_MNC  
Gapop 60.0 , Gapext 60.0

Searched: 442118 seqs, 280819700 residues

Word size : 25

Total number of hits satisfying chosen parameters: 12207

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	898	100.0	898	10	US-09-832-320-1
2	612	68.2	612	10	US-09-832-320-3
3	43	4.8	447	10	US-09-770-444-759
4	43	4.8	1077	9	US-09-764-868-497
5	42	4.7	111	10	US-09-925-299-635
6	42	4.7	365	10	US-09-960-352-5456
7	42	4.7	495	10	US-09-864-761-1723
8	42	4.7	595	10	US-09-867-552A-31
9	42	4.7	707	10	US-09-764-864-19
10	42	4.7	938	10	US-09-037-657-16
11	42	4.7	1041	10	US-09-925-300-440
12	42	4.7	1119	9	US-09-796-753-133
13	42	4.7	1267	10	US-09-925-301-188
14	42	4.7	1300	10	US-09-822-849A-3
15	42	4.7	1431	10	US-09-764-864-695
16	42	4.7	1436	10	US-09-764-864-272
17	42	4.7	1659	9	US-09-796-692-667
18	42	4.7	1686	10	US-09-925-301-157
19	42	4.7	1711	10	US-09-935-390A-17

20	42	4.7	1824	10	US-09-814-786-1	Sequence 1, Appl1
21	42	4.7	1877	12	US-10-007-399-2	Sequence 2, Appl1
22	42	4.7	2104	10	US-09-925-300-288	Sequence 288, App
23	42	4.7	2519	12	US-10-139-262-9	Sequence 9, Appl1
24	42	4.7	3084	10	US-09-764-864-89	Sequence 89, Appl1
25	42	4.7	3397	10	US-09-925-301-414	Sequence 414, App
26	42	4.7	5354	9	US-09-815-923-1	Sequence 1, Appl1
27	42	4.6	112	9	US-09-768-827-51	Sequence 51, Appl1
28	41	4.6	185	9	US-10-046-935-1615	Sequence 1615, Ap
29	41	4.6	185	9	US-09-878-178-1615	Sequence 1615, Ap
30	41	4.6	202	10	US-09-925-299-573	Sequence 573, App
31	41	4.6	211	10	US-09-867-701-2599	Sequence 2599, Ap
32	41	4.6	217	10	US-09-878-574-12266	Sequence 12266, Ap
33	41	4.6	255	10	US-09-878-574-12064	Sequence 13064, A
34	41	4.6	276	10	US-09-878-574-14211	Sequence 14211, A
35	41	4.6	292	10	US-09-960-352-12176	Sequence 12176, A
36	41	4.6	302	10	US-09-925-299-710	Sequence 710, App
37	41	4.6	308	9	US-09-954-531-79	Sequence 79, Appl1
38	41	4.6	317	10	US-09-954-531-263	Sequence 263, App
39	41	4.6	317	9	US-09-960-352-3366	Sequence 3366, Ap
40	41	4.6	376	9	US-09-933-797-343	Sequence 343, App
41	41	4.6	383	10	US-09-960-352-5366	Sequence 5366, Ap
42	41	4.6	418	10	US-09-960-352-10650	Sequence 10650, A
43	41	4.6	444	10	US-09-960-352-1281	Sequence 1281, Ap
44	41	4.6	456	10	US-09-960-352-9659	Sequence 9659, Ap
45	41	4.6	523	10	US-09-962-436-68	Sequence 68, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-832-320-1
; Sequence 1, Application US/09832320
; Patent No. US20010049834A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Edmund H.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Pathogenesis-Related
; FILE REFERENCE: 35718/214291
; CURRENT APPLICATION NUMBER: US/09/832,320
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/195,801
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)...(674)
US-09-832-320-1
Query Match      100.0%; Score 898; DB 10; Length 898;
Best Local Similarity 100.0%; Pred. No. 8.1e-299;
Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CTGCAGCAGCTCGACGCTCATCTGAGCCATTACTAGATCAACCACTCCAGATCT 60
    |||||||
DB 1 CTGCAGCAGCTCGACGCTCATCTGAGCCATTACTAGATCAACCACTCCAGATCT 60
OY 61 CATGCGGACCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
    |||||||
DB 61 CATGCGGACCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
OY 121 CATTGCTTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
    |||||||
DB 121 CATTGCTTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
OY 181 CCGGCTCTCTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
    |||||||
```

```
|||||
Db 181 CCGGGCTCTATGCGGGGCGGGGGGGGGGGGACCAAGGCCACAGGTGGCCACC 240
Qy 241 GCAGCGAGCAAGCGAGCGAGCGAGTACTGCGCGCCGACACAGGAGCGCGG 300
Db 241 GCAGCGAGCAAGCGAGCGAGCGAGTACTGCGCGCCGACACAGGAGCGCGG 300
Qy 301 CCGTGGGCTGGCCCGCTGCGGTGGAAACGGGGCTTGGCTTGGCGCGCGCGG 360
Db 301 CCGTGGGCTGGCCCGCTGCGGTGGAAACGGGGCTTGGCTTGGCGCGCGG 360
Qy 361 TGGCGACAGCGCGGAGGAGGGGAGTCCGCTTGGCGAGCGAGTGGCGGACG 420
Db 361 TGGCGACAGCGCGGAGGAGGGGAGTCCGCTTGGCGAGCGAGTGGCGGACG 420
Qy 421 GCGGCAACCAAGGGGTGGGCGAGTACCGCGCGCGCCCGCGAGTGGTGGG 480
Db 421 GCGGCAACCAAGGGGTGGGCGAGTACCGCGCGCGCCCGCGAGTGGTGGG 480
Qy 481 TGGCGAGAGGGGGTACTACACCCGACCAACAGTGGCGCGCGGGGGGAGTGG 540
Db 481 TGGCGAGAGGGGGTACTACACCCGACCAACAGTGGCGCGCGGGGGGAGTGG 540
Qy 541 GCACGTACACGACGAGTGGTGGCGCAACACCGCGAGTGGCGCGAGCCAGCT 600
Db 541 GCACGTACACGACGAGTGGTGGCGCAACACCGCGAGTGGCGCGAGCCAGCT 600
Qy 601 GCGGCAAGGGGGCGACGCTGCAAGCTGCTGTAACACCGGAGGAGTGGAGG 660
Db 601 GCGGCAAGGGGGCGACGCTGCAAGCTGCTGTAACACCGGAGGAGTGGAGG 660
Qy 661 AGAGCCCTACTAGCTAGCTAGCTATCAGTGTAGCGAGCGCAACGTCGCCG 720
Db 661 AGAGCCCTACTAGCTAGCTAGCTATCAGTGTAGCGAGCGCAACGTCGCCG 720
Qy 721 CCGGGGAGGAGGAGTACTAGTGTTCATCAGTCTTCTAGTTGGGCAAGGAG 780
Db 721 CCGGGGAGGAGGAGTACTAGTGTTCATCAGTCTTCTAGTTGGGCAAGGAG 780
Qy 781 TTTTGTGTGATCGGTGTGTTCTTGTGTTGTGACAACTGTTGTTGTGTA 840
Db 781 TTTTGTGTGATCGGTGTGTTCTTGTGTTGTGACAACTGTTGTTGTGTA 840
Qy 841 TCAGCTTTTGTGCGGGTAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG 898
Db 841 TCAGCTTTTGTGCGGGTAAAAAAGGGGGGGGGGGGGGGGGGGGGGGGGGG 898
```

```
RESULT 2
US-09-832-320-3
; Sequence 3, Application US/09832320
; Patent No. US20010049834A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: Crane, Edmund H.
```

```
APPLICANT: Crane, Virginia C.
```

```
TITLE OF INVENTION: Maize Pathogenesis-Related
```

```
TITLE OF INVENTION: Polynucleotide and Methods of use
```

```
FILE REFERENCE: 35718/214291
```

```
CURRENT APPLICATION NUMBER: US/09/832,320
```

```
PRIOR FILING DATE: 2001-04-10
```

```
PRIOR APPLICATION NUMBER: US 60/195,801
```

```
NUMBER OF SEQ ID NOS: 3
```

```
SOFTWARE: FastSeq for Windows Version 4.0
```

```
SEQ ID NO 3
```

```
LENGTH: 612
```

```
TYPE: DNA
```

```
ORGANISM: Zea mays
```

```
FEATURE:
```

```
NAME/KEY: CDS
```

```
LOCATION: (1)...(612)
```

```
US-09-832-320-3
```

```
Query Match 68.2%; Score 612; DB 10; Length 612;
Best Local Similarity 100.0%; Pred. No. 5,1e-201;
Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 ATGGGCGATCGGCGAGCGACACCACTCTCTCTGCTGCGCGCGGCGGCGGCGG 122
Db 1 ATGGGCGATCGGCGAGCGACCACTCTCTCTCTGCTGCGCGCGGCGGCGGCGG 60
Qy 123 TCGTTGCTCTGCGACCCCTCTGCGCTGTGCGCGCGCGCGCGCGCGCGCGG 182
Db 61 TCGTTGCTCTGCGACCCCTCTGCGCTGTGCGCGCGCGCGCGCGCGCGCGG 120
Qy 183 CCGGCTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
Db 121 CCGGCTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy 243 AGCGGCAACAGCGGAGCGGCGGAGTACTGCGCGCGCGCAACAGGAGCGCGCGG 302
Db 181 AGCGGCAACAGCGGAGCGGCGGAGTACTGCGCGCGCGCAACAGGAGCGCGG 240
Qy 303 GTGGGCGTGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362
Db 241 GTGGGCGTGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Qy 363 GCGGAGCGCGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422
Db 301 GCGGAGCGAGCGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Qy 423 GCGAACAAGGGGTGGGCGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGG 482
Db 361 GCGAACAAGGGGTGGGCGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGG 420
Qy 483 GCGGAGGGGCGTACTACACCAACCAACAGTGTGCGCGCGCGCGCGCGCGG 542
Db 421 GCGGAGGGGCGTACTACACCAACCAACAGTGTGCGCGCGCGCGCGCGCGG 480
Qy 543 ACGTACACGCGAGTGTGTGGCGCAACCGCGGAGTGGGTGGCGGCGGCGGCGG 602
Db 481 ACGTACACGCGAGTGTGTGGCGCAACCGCGGAGTGGGTGGCGGCGGCGGCGG 540
Qy 603 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 662
Db 541 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Qy 663 AGCCCTACTAG 674
Db 601 AGCCCTACTAG 612
```

```
RESULT 3
US-09-770-444-759/C
; Sequence 759, Application US/09770444
; Patent No. US20020023280A1
```

```
GENERAL INFORMATION:
```

```
APPLICANT: Goriach, Jörn
```

```
APPLICANT: An, Yong-Qiang
```

```
APPLICANT: Hamilton, Carol M.
```

```
APPLICANT: Price, Jennifer L.
```

```
APPLICANT: Raines, Tracy M.
```

```
APPLICANT: Yu, Yang
```

```
APPLICANT: Rameaka, Joshua G.
```

```
APPLICANT: Page, Amy
```

```
APPLICANT: Matthew, Abraham V.
```

```
APPLICANT: Ledford, Brooke L.
```

```
APPLICANT: Moessner, Jeffrey P.
```

```
APPLICANT: Haas, William David
```

```
APPLICANT: Garcia, Carlos A.
```

```
APPLICANT: Krieker, Maya
```

```
APPLICANT: Slader, Ted
```

```
APPLICANT: Davis, Keith R.
```

```
APPLICANT: Allen, Keith
```

```
APPLICANT: Hoffman, Neil
```

```
APPLICANT: Hurban, Patrick
```

```

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: Chailana
; CURRENT APPLICATION NUMBER: US/09/770,444
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 759
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-759

```

```

Query Match
Best Local Similarity 100.0%; Score 43; DB 10; Length 447;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 74 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 32

```

```

RESULT 4
US-09-764-868-497
; Sequence 497, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT32
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 497
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-497

```

```

Query Match
Best Local Similarity 100.0%; Score 43; DB 9; Length 1077;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 1026 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1068

```

```

RESULT 5
US-09-925-299-635
; Sequence 635, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 635
; LENGTH: 111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (19)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (24)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (35)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (38)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (109)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-635

```

```

Query Match
Best Local Similarity 100.0%; Score 42; DB 10; Length 111;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 856 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 897
Db 58 GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 99

```

```

RESULT 6
US-09-960-352-5456
; Sequence 5456, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; PRIOR FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5456
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 24-BOWMS1-009-Q1-E1-F7
US-09-960-352-5456

```

```

Query Match
Best Local Similarity 100.0%; Score 42; DB 10; Length 365;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 18 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 59

```

```

RESULT 7
US-09-864-761-1723
; Sequence 1723, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761

```

```

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1723
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011347.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
US-09-864-761-1723

Query Match 4.7%; Score 42; DB 10; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 234 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 275
```

```

RESULT 8
US-09-967-552A-31
; Sequence 31, Application US/09967552A
; Patent No. US20020124282A1
; GENERAL INFORMATION:
; APPLICANT: Danilevskaya, Olga
; APPLICANT: Butler, Karlene H.
```

```

; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Sakai, Hajime
; APPLICANT: Simmons, Carl R.
; APPLICANT: Wang, Zude
; APPLICANT: Famodu, Omojayo
; APPLICANT: Hantke, Sabine
; TITLE OF INVENTION: Plant Reproduction Polynucleotides and Methods of Use
; FILE REFERENCE: DD0010R
; CURRENT APPLICATION NUMBER: US/09/967,552A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/US00/23735
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: US 60/151,575
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(360)
US-09-967-552A-31
```

```

Query Match 4.7%; Score 42; DB 10; Length 595;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 525 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 566
```

```

RESULT 9
US-09-764-864-19
; Sequence 19, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-19
```

```

Query Match 4.7%; Score 42; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 634 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 675
```

```

RESULT 10
US-09-037-657-16
; Sequence 16, Application US/09037657A
; Patent No. US20020045741A1
; GENERAL INFORMATION:
; APPLICANT: Hilton, Douglas J.
; APPLICANT: Nicola, Nicos A.
; APPLICANT: Farley, Allison
; APPLICANT: Wilson, Tracy
; APPLICANT: Zhang, Jian-Guo
```



```

; APPLICANT: Alexander, Warren
; APPLICANT: Rakar, Steven
; APPLICANT: Fabri, Louis
; APPLICANT: Kojima, Tetsuo
; APPLICANT: Maeda, Masatsugu
; APPLICANT: Kikuchi, Yasufumi
; APPLICANT: Nash, Andrew
; TITLE OF INVENTION: A NOVEL HAEMPOIETIN RECEPTOR AND GENETIC SEQUENCES
; FILE REFERENCE: DAVIES COLLISON CAVE (CIP)
; CURRENT APPLICATION NUMBER: US/09/037,657A
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 08/928,720
; EARLIER FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(465)
; OTHER INFORMATION: Description of Unknown Organism: Murine NR6.3
US-09-037-657-16

Query Match
Best Local Similarity 100.0%; Score 42; DB 10; Length 938;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 884 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 925
```

```

RESULT 11
US-09-925-300-440
; Sequence 440, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 440
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1025)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1030)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1039)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-440
```

```

Query Match
Best Local Similarity 100.0%; Score 42; DB 10; Length 1041;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
Db 979 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1020
```

```

RESULT 12
US-09-796-753-133
; Sequence 133, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 08/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 133
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-753-133
```

```
Query Match          4.7%; Score 42; DB 9; Length 1119;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
DB 1040 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1081

RESULT 13
US-09-925-301-188
; Sequence 188, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05862
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 188
; LENGTH: 1267
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-188

Query Match          4.7%; Score 42; DB 10; Length 1267;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
DB 1195 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1236

RESULT 14
US-09-822-849A-3
; Sequence 3, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakari
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1300
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-822-849A-3

Query Match          4.7%; Score 42; DB 10; Length 1300;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
DB 1232 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1273

RESULT 15
US-09-764-864-695
; Sequence 695, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; PRIOR application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 695
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-695

Query Match          4.7%; Score 42; DB 10; Length 1431;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 857 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898
DB 1390 GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1431

Search completed: February 22, 2003, 06:10:54
Job time : 100.611 secs
```

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 04:05:06 ; Search time 985.279 Seconds

(without alignments)  
10059,722 Million cell updates/sec

Title: US-09-832-320-3

Perfect score: 612  
Sequence: 1 atggcgacacgcgcgcagcca.....aggcgccagagccctactag 612

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 25

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: em\_estfun:\*  
15: em\_estom:\*  
16: em\_gss:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	12.7	297	10	AW923743 DGI_59_A0
2	75	12.3	486	10	BE512404 BE512404
3	50	8.2	362	13	BM101325 BM101325
4	50	8.2	996	12	BG343299 BG343299
5	40	6.5	681	9	AU082529 AU082529
6	35	5.7	293	9	AU029886 AU029886

# ALIGNMENTS

Result	Score	Query Match	Length	ID	Description
1	78	12.7	297	10	AW923743 DGI_59_A0
2	75	12.3	486	10	BE512404 BE512404
3	50	8.2	362	13	BM101325 BM101325
4	50	8.2	996	12	BG343299 BG343299
5	40	6.5	681	9	AU082529 AU082529
6	35	5.7	293	9	AU029886 AU029886

FEATURES

source

1. 297

/organism="Sorghum bicolor"

/db\_xref="taxon:4558"

/clone\_lib="Dark grown 1 (DGI)"

/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda Zap; Site\_1: XhoI; Site\_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

BASE COUNT

46 a 86 c 97 g 67 t 1 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 2.2e-23; Length 297;

Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 ACCTGGCCACGCGGCGCCACGCTGCTGCTGACACCGCCGCAACGTCAG 594

DB 43 ACTGGCCACGCGGCGCCACGCTGCTGCTGACACCGCCGCAACGTCAG 102

QY 595 GGCACAGCGCCCTACTAG 612

DB 103 GGCACAGCGCCCTACTAG 120

RESULT 2

LOCUS BE512404 486 bp mRNA linear EST 07-AUG-2000

DEFINITION 946070F05.y1 946 - tassal primordium prepared by Schmidt lab Zea

ACCESSION BE512404

VERSION BE512404.1 GI:9733652

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays



TITLE	Rice cDNA from panicle (2000)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-0602, Japan Tel: 81-298-38-7441 Fax: 81-298-38-7468 Email: tsasaki@abrc.affrc.go.jp/ PROJECT="RGP"
FEATURES	Location/Qualifiers
SOURCE	1..681 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nippondare" /db_xref="taxon:39947" /clone="E30820" /clone_1lb="Rice panicle shorter than 3cm" /dev_stage="shorter than 3cm" /note="Organ: panicle"
BASE COUNT	136 a 187 c 214 g 141 t 3 others
ORIGIN	

MODEL	BE125871	288 bp	mRNA	linear	EST 19-JUL-2000
LOCUS	DGL_59_A04	b1_A002	Dark Green 1	(DGL)	Sorghum bicolor cDNA, mRNA
DEFINITION	sequence.				

VERSION	KEYWORDS
BE125871.1	GI:8548461
SOURCE	ORGANISM
EST.	
Sorghum	
Sorghum bicolor	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC	
Clade; Panicoideae; Andropogoneae; Sorghum.	

REFERENCE 1 (bases 1 to 288)  
 AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt  
 TITLE Ap EST database from Sorghum: dark-grown seedlings  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmp@prattuga.edu

Sequences have been trimmed to exclude polyA, vector and regions  
 below phred quality 16. The threshold for highest quality sequence  
 is 20.

Seq primer: JEN REV  
 High quality sequence stop: 285  
 POLYA-No.

FEATURES  
 SOURCE location/Qualifiers

1..288  
 /organism="Sorghum bicolor"  
 /db\_xref="taxon:4558"  
 /clone\_lib="Dark Grown 1 (DGI)"  
 /note="Organ: 5-day-old dark-grown seedlings; Vector:  
 lambda zap; Site.1: XhoI; Site.2: EcoRI; The library was  
 made from poly-A RNA in the cloning vector lambda ZAP II.  
 Clones to be sequenced were prepared by mass excision."  
 46 a 123 c 76 g 42 t 1 others

BASE COUNT  
 ORIGIN

Query Match 5.1%; Score 31; DB 10; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 0.0039;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 CCTCTCGCGCTCTGCGCGCGCGCGCGCG 108  
 ||||||||||||||||||||||||||||||||  
 DB 193 CCTCTCGCGCTCTGCGCGCGCGCGCGCG 223

Search completed: February 22, 2003, 06:06:24  
 Job time : 988.279 secs

Qy	361	GCACAAACGAGGGTGGGCGCAAGCTAACCGCGCGCGCCCGCGAGAGTGTGGCTTGGGTG	420
Db	361	GCACAAACGAGGGTGGGCGCAAGCTAACCGCGCGCGCCCGCGAGAGTGTGGCTTGGGTG	420
Qy	421	GGCGAGGGGCGGTACTATACACCCACGCCAACAACAGTGGCGCGGGGCGGCACTGGCGC	480
Db	421	GGCGAGGGGCGGTACTATACACCCACGCCAACAACAGTGGCGCGGGGCGGCACTGGCGC	480
Qy	481	ACGTAAACGACAGGTGTGTGGCGCAACACCGCGAGTGGGTGGCGCAGGCCAAGTGC	540
Db	481	ACGTAAACGACAGGTGTGTGGCGCAACACCGCGAGTGGGTGGCGCAGGCCAAGTGC	540
Qy	541	GCCACGGGCGCCACGCTACACCTCTGGCTTAAACAACCGCACGGCAACGTGCGAGGCCAG	600
Db	541	GCCACGGGCGCCACGCTACACCTCTGGCTTAAACAACCGCACGGCAACGTGCGAGGCCAG	600
Qy	601	AGCCCCCTACTAG 612	
Db	601	AGCCCCCTACTAG 612	

## RESULT 2

ID	ABA96417	standard; cDNA; 898 BP.
...		

AC ABA96417;

DT 02-APR-2002 (first entry)

DE Malze PR1-C10 encoding cDNA SEQ ID NO 1.

KW Maize; pathogen-related; PR1-C10; plant; transgenic; gene; ss.

**Zea mays.**

FH	Key	Location/Qualifiers
ET	CDS	63 674

```

ENT      /*tag= a
ENT      /product= "BP1-c10"

```

XX  
PN  
HS2001049834-A1

06-DEC-2001  
XX  
PD

AA 10-APR-2001: 2001US-0832320

AA  
PR 10-APR-2000: 2000US-195801P.

PA (CRAN/) CRANE E H.

XX

XX

DR P-PSDB; AAM48742.

PT New nucleic acid encoding a pathogen-related protein isolated from

PT enhanced disease resistance

PS Claim 1; Page 30-31; 34pp; English.

CC The invention relates to an isolated nucleic acid encoding a

CC plants for enhanced disease resistance.

sequence 898 bp; 161 A; 295 C; 309 G; 133 T; 0 other;

Query Match	100.0%;	Score 612;	DB 24;	Length 898;
Best local similarity	100.0%;	Pred. No. 2.1e-250;		

```

    matches 0; conservative 0; mismatches 0; indels 0; gaps 0;

```

Db	63	ATTGGCGACTCGCGGACGACACACACACTCTCTCTCTCTCCCGCGCCATGGCCACGGCG	122
Oy	61	TGCTTGCTCTCTGGCCACACCTCTCTGGGCTTGGCGCGCGCGCGCCACCCACGCGCG	120
Db	123	TGCTTGCTCTCTGGCCACACCTCTCTGGCGCTTGGCGCGCGCGCGCCACCCACGCGCG	182
Oy	121	CGGTCCTCATMGCCGAGCGCGCGCGCGCGGCGGTCACCAAGGCGCACAGGTCGGACCGGC	180
Db	183	CGGTCCTCATMGCCGAGCGCGCGCGCGCGGCGGTCACCAAGGCGCACAGGTCGGACCGGC	242
Oy	181	AGCGGACGACACGCGACGCGGAGTACTTGGCGCGCCACAACTAGCGCGCGCGCG	240
Db	243	AGCGGACGACACGCGACGCGGAGTACTTGGCGCGCCACAACTAGCGCGCGCGCG	302
Oy	241	GTGGGCGTGGGCCCGCTGGGGTGGACGCGGGGCTTGCGCGCGCGCGCGGGACGGTG	300
Db	303	GTGGGCGTGGGCCCGCTGGGGTGGACGCGGGGCTTGCGCGCGCGCGCGGGACGGTG	362
Oy	301	GCGCAGCAGCGCGCGAGGGCGGGTGGCGTTCGCGGACGTGGGGGCCAGCCCTTACGGC	360
Db	363	GCGCAGCAGCGCGCGAGGGCGGGTGGCGTTCGCGGACGTGGGGGCCAGCCCTTACGGC	422
Oy	361	GCGCAGCAGGGGGTGGGCGAGCTACCGCGCGCGCCCGCGGAGTGGTGCGCTTGGGTG	420
Db	423	GCGCAGCAGGGGGTGGGCGAGCTACCGCGCGCGCCCGCGGAGTGGTGCGCTTGGGTG	482
Oy	421	GCGCAGGCGGGGTACTACACCCAGCCACAACTGGCGCGCGGGCGGGCGAGTGGCGC	480
Db	483	GCGCAGGCGGGGTACTACACCCAGCCACAACTGGCGCGCGGGCGGGCGAGTGGCGC	542
Oy	481	ACGTACACGAGGGTGTGGGGCGCAACCGCGAGGTGGGGTGGCGCDAGGCCACGTGC	540
Db	543	ACGTACACGAGGGTGTGGGGCGCAACCGCGAGGTGGGGTGGCGCDAGGCCACGTGC	602
Oy	541	GCCACGCGGCCACGCTACCGCTTGCTGTACAAACCCGACGCGACAGTGCAGGGCCAG	600
Db	603	GCCACGCGGCCACGCTACCGCTTGCTGTACAAACCCGACGCGACAGTGCAGGGCCAG	662
Oy	601	AGCCCGCTACTAG 612	
Db	663	AGCCCGCTACTAG 674	

Search completed: February 22, 2003, 04:15:14  
Job time : 158.877 secs

ID	ABA96418 standard; cDNA; 612 BP.
XX	
AC	ABA96418;
XX	
DT	02-APR-2002 (first entry)
XX	
DE	Maize PR1-C10 coding sequence SEQ ID NO 3.
XX	
KW	Maize; pathogen-related; PR1-C10; plant; transgenic; gene; ss.
XX	
OS	Zea mays.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..612
FT	/tag= a
FT	/product= "PR1-C10"
XX	
PN	US2001049834-A1.
XX	
PD	06-DEC-2001.
XX	
PF	10-APR-2001; 2001US-0832320.
XX	
PR	10-APR-2000; 2000US-195801P.
XX	
PA	(CRAN/) CRANE E H.
XX	(CRAN/) CRANE V C.
PI	Crane EH, Crane VC;
XX	
DR	WPI; 2002-121407/16.
XX	P-PSDB; AAM48742.
PT	New nucleic acid encoding a pathogen-related protein isolated from
PT	maize and designated PR1-C10, useful for transforming plants for
PT	enhanced disease resistance -
XX	
PS	Claim 1; Page 32; 34pp; English.
XX	
CC	The invention relates to an isolated nucleic acid encoding a
CC	pathogen-related protein PR1-C10. The nucleic acid is used to transform
CC	plants for enhanced disease resistance.
XX	
SQ	Sequence 612 BP; 80 A; 231 C; 239 G; 62 T; 0 other;
	Query Match 100.0%; Score 612; DB 24; Length 612;
	Best Local Similarity 100.0%; Pred. No. 2,3e+250;
	Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY	1 ATGGCGCAGTGCAGGCAACCACACACACTTCCTCCGCCGCCCATGGCCACGGC 60
DB	1 ATGGCGCAGTGCAGGCAACCACACACACTTCCTCCGCCGCCCATGGCCACGGC 60
QY	61 TGCATTGCTCTCGCCACCCCTCTCGGCTGTGGCCGCGCGCGCGCGCACGGCGG 120
DB	61 TGCATTGCTCTCGCCACCCCTCTCGGCTGTGGCCGCGCGCGCGCGCACGGCGG 120
QY	121 CGGGTCTCTATGCGCGGCGGCGCGGCGCGGTACCAAGGCGAGAGGGTGACCGGC 180
DB	121 CGGGTCTCTATGCGCGGCGGCGGCGGCGGTACCAAGGCGAGAGGGTGACCGGC 180
QY	181 AGCGGCGAGCAACGCGAGCGGCGAGAGTAGTACTGTGGCGCGGCACAAACGAGCGCGCGGCG 240
DB	181 AGCGGCGAGCAACGCGAGCGGCGAGAGTAGTACTGTGGCGCGGCACAAACGAGCGCGCGGCG 240
QY	241 GTGGGCGTGGCCCCGCTGCGGTTGGAAGCGGCGCTGGCTTCGGGGCGCGGGGACGGTG 300
DB	241 GTGGGCGTGGCCCCGCTGCGGTTGGAAGCGGCGCTGGCTTCGGGGCGCGGGGACGGTG 300
QY	301 GCGGAGAGGCGGCGGAGGGGCGGTTGCGCTGCGGAGCTGGGGGCGCAGCCCTACGGC 360
DB	301 GCGGAGAGGCGGCGGAGGGGCGGTTGCGCTGCGGAGCTGGGGGCGCAGCCCTACGGC 360



GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 04:08:26 ; Search time 60.3894 Seconds  
(without alignments)  
5691.782 Million cell updates/sec

Title: US-09-832-320-3

Perfect score: 612  
Sequence: 1 atggcgacatcgccgagccca.....aggagccagagccctactag 612

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 44218 seqs, 280819700 residues

Word size: 25

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database: Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	612	100.0	612	10	US-09-832-320-3
2	612	100.0	898	10	US-09-832-320-1

## ALIGNMENTS

RESULT 1  
US-09-832-320-3  
; Sequence 3, Application US/09832320  
; Patent No. US20010049834A1  
; GENERAL INFORMATION:  
; APPLICANT: Crane, Edmund H.  
; APPLICANT: Crane, Virginia C.  
; TITLE OF INVENTION: Maize Pathogenesis-Related  
; FILE REFERENCE: 35718/214291  
; CURRENT APPLICATION NUMBER: US/09/832,320  
; CURRENT FILING DATE: 2001-04-10

PRIOR APPLICATION NUMBER: US 60/195,801  
; PRIOR FILING DATE: 2000-04-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(612)  
US-09-832-320-3

Query Match 100.0%; Score 612; DB 10; Length 612;  
Best Local Similarity 100.0%; Pred. No. 2.9e+266;  
Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCGACATCGCGAGCAGCACACACCTCTGCTCCCGCGCCATGAGCCAGCG	60
DB	1	ATGGCGACATCGCGAGCAGCACACACCTCTGCTCCCGCGCCATGAGCCAGCG	60
QY	61	TGCTTGTCTTCTTCTGCGACCTCTGCGCGCGCGCGCGCGCGCGCGCG	120
DB	61	TGCTTGTCTTCTGCGACCTCTGCGCGCGCGCGCGCGCGCGCGCGCG	120
QY	121	CGCTCTCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180
DB	121	CGCTCTCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180
QY	181	AGCGGACAGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	240
DB	181	AGCGGACAGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	240
QY	241	GTCG	300
DB	241	GTCG	300
QY	301	GCGGCG	360
DB	301	GCGGCG	360
QY	361	GCGGCG	420
DB	361	GCGGCG	420
QY	421	GCGGCG	480
DB	421	GCGGCG	480
QY	481	ACGTACAGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	540
DB	481	ACGTACAGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	540
QY	541	GCGGCG	600
DB	541	GCGGCG	600
QY	601	AGCCCTACTAG 612	
DB	601	AGCCCTACTAG 612	

RESULT 2  
US-09-832-320-1  
; Sequence 1, Application US/09832320  
; Patent No. US20010049834A1  
; GENERAL INFORMATION:  
; APPLICANT: Crane, Edmund H.  
; APPLICANT: Crane, Virginia C.  
; TITLE OF INVENTION: Maize Pathogenesis-Related  
; FILE REFERENCE: 35718/214291  
; CURRENT APPLICATION NUMBER: US/09/832,320

```

; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/195,801
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ. ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(674)
US-09-832-320-1

```

```

Query Match      100.0%; Score 612; DB 10; Length 898;
Best Local Similarity 100.0%; Pred. No. 2,8e-266;
Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ATGGGCACTGCGCCACACACACCTCTGCTCCCGCGCCCATGGCCACGGCG 60
   |||||||
Db 63 ATGGGCACTGCGCCACACACACCTCTGCTCCCGCGCCCATGGCCACGGCG 122
QY 61 TGCCTGCTCTGCGCCACCTCTGCTGCGCGCGCGCGCCGACCCACGGCG 120
   |||||||
Db 123 TGCCTGCTCTGCGCCACCTCTGCTGCGCGCGCGCGCCGACCCACGGCG 182
QY 121 GCGTCTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
   |||||||
Db 183 GCGTCTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242
QY 181 AGCGGCAACAGCGAGCGGCGAGTACCTGCGCGCGCGCGCGCGCGCG 240
   |||||||
Db 243 AGCGGCAACAGCGAGCGGCGAGTACCTGCGCGCGCGCGCGCGCGCG 302
QY 241 GTGGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
   |||||||
Db 303 GTGGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 362
QY 301 GCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
   |||||||
Db 363 GCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 422
QY 361 GCGAACGAGGGGTGGGAGGTACCGCGCGCGCGCGCGCGCGCGCG 420
   |||||||
Db 423 GCGAACGAGGGGTGGGAGGTACCGCGCGCGCGCGCGCGCGCGCG 482
QY 421 GCGGAGGCGGTACTACACCCACCAACACGTCGCCCGCGCGCGCG 480
   |||||||
Db 483 GCGGAGGCGGTACTACACCCACCAACACGTCGCCCGCGCGCGCG 542
QY 481 ACGTACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
   |||||||
Db 543 ACGTACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
QY 541 GCCACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
   |||||||
Db 603 GCCACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 662
QY 601 AGCCCTTACTAG 612
   |||||||
Db 663 AGCCCTTACTAG 674

```

Search completed: February 22, 2003, 06:10:55  
Job time : 61.3894 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 22, 2003, 03:01:40 ; Search time 77 Seconds

(Without alignments)  
808,511 Million cell updates/sec

Title: US-09-832-320-2

Sequence: 1 MAHSRHHHLLLPAPMATA.....ATLTCLYPNHGNVQGSPPY 203

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlp  
-O/cgn2\_1/USPTO.spool/US09832320/rumat.20022003.092738.14818/app.query.fasta\_1.391  
-Db-Issued Patents.NA -OPMT=fastap -SUFFIX=g2n.rn1 -MINMATCH=0.1 -LOOPT=0  
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdl  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPRN=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=0509832320.ecgn\_1.1.61.rumat.20022003.092738.14818 -MCP=6 -ICP=3  
-NO\_XLPEXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEDOUT=120  
-WARN.TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	333	30.7	749	4	US-09-257-583-12
2	304	28.0	771	1	US-08-181-271A-10
3	304	28.0	771	1	US-08-449-315-10
4	304	28.0	771	1	US-08-444-803-10
5	304	28.0	771	1	US-08-449-043-10
6	304	28.0	771	1	US-08-456-265A-10
7	304	28.0	771	1	US-08-455-416-10
8	304	28.0	771	1	US-08-455-244-10
9	304	28.0	771	1	US-08-454-876-10
10	304	28.0	771	2	US-08-457-364-10
11	304	28.0	771	2	US-08-456-262-10
12	304	28.0	771	2	US-08-456-240-10

13	304	28.0	771	2	US-08-455-736-10	Sequence 10, Appl
14	304	28.0	771	4	US-09-350-600-10	Sequence 10, Appl
15	304	28.0	771	4	US-08-181-271A-11	Sequence 11, Appl
16	286	27.3	696	1	US-08-449-315-11	Sequence 11, Appl
17	286	27.3	696	1	US-08-444-803-11	Sequence 11, Appl
18	286	27.3	696	1	US-08-449-043-11	Sequence 11, Appl
19	286	27.3	696	1	US-08-456-265A-11	Sequence 11, Appl
20	286	27.3	696	1	US-08-455-416-11	Sequence 11, Appl
21	286	27.3	696	1	US-08-455-244-11	Sequence 11, Appl
22	286	27.3	696	1	US-08-454-876-11	Sequence 11, Appl
23	286	27.3	696	2	US-08-457-364-11	Sequence 11, Appl
24	286	27.3	696	2	US-08-456-262-11	Sequence 11, Appl
25	286	27.3	696	2	US-08-455-736-11	Sequence 11, Appl
26	286	27.3	696	2	US-08-457-364-11	Sequence 11, Appl
27	286	27.3	696	2	US-08-455-244-11	Sequence 11, Appl
28	286	27.3	696	2	US-08-454-876-11	Sequence 11, Appl
29	286	27.3	696	4	US-09-350-600-11	Sequence 11, Appl
30	286	27.3	696	4	US-09-257-583-6	Sequence 6, Appl
31	286	26.5	860	1	US-08-181-271A-33	Sequence 33, Appl
32	286	26.5	860	1	US-08-449-315-33	Sequence 33, Appl
33	286	26.5	860	1	US-08-444-803-33	Sequence 33, Appl
34	286	26.5	860	1	US-08-449-043-33	Sequence 33, Appl
35	286	26.5	860	1	US-08-456-265A-33	Sequence 33, Appl
36	286	26.5	860	1	US-08-455-416-33	Sequence 33, Appl
37	286	26.5	860	1	US-08-455-244-33	Sequence 33, Appl
38	286	26.5	860	1	US-08-454-876-33	Sequence 33, Appl
39	286	26.5	860	2	US-08-457-364-33	Sequence 33, Appl
40	286	26.5	860	2	US-08-456-262-33	Sequence 33, Appl
41	286	26.5	860	2	US-08-456-240-33	Sequence 33, Appl
42	286	26.5	860	2	US-08-455-736-33	Sequence 33, Appl
43	286	26.5	860	2	US-08-457-364-33	Sequence 33, Appl
44	286	26.5	860	4	US-09-350-600-33	Sequence 9, Appl
45	286	26.4	809	1	US-08-181-271A-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-09-257-583-12  
; Sequence 12, Application US/09257583A  
; Patent No. 6429362  
; GENERAL INFORMATION:  
; APPLICANT: Crane, Virginia  
; TITLE OF INVENTION: Family of Maize PR-1 Genes And Promoters  
; FILE REFERENCE: 5718-32, 035718/175219  
; CURRENT APPLICATION NUMBER: US/09/257,583A  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 749  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (92)..(583)  
US-09-257-583-12

## Alignment Scores:

Pred. No.: 1,14e-19  
Score: 333.00  
Percent Similarity: 48.13%  
Best Local Similarity: 37.65%  
Query Match: 30.69%  
DB: 4  
Gaps: 8  
US-09-832-320-2 (1-203) x US-09-257-583-12 (1-749)

5 ArgSerHisHisIsteuIeuIeuIeuProAla-----  
|||||||  
26 CGATTCACATGAGCTTCGACTGCTGCTCATATACAGTTCACAGCAACAA 85  
16 ---Prometala-----ThrAlaCysIeuIeuIeuAlaThrIeuIeuAlaLeuCysAla 32

DB 86 CAACACATGCGACCGAGCTAGCTGCTCCCTAGCTGCGCCATGCGCCATGCTGCTG 145  
OY 33 AAlaProAlaProThrisglYalaarYalLeuMetProglYglYalaYalaValThr 52  
DB 146 GCGCGCTGCG----- 154  
OY 53 LysAlaInglInglYglYThrglySerGlySerAsnAlaThrAlaAspGluTyrLeuAla 72  
DB 155 -----ACGCCCGCAACTGCGCGCAGAC---TAGGTGAC 167  
OY 73 ProHisanglAlaArgAlaAlaValGlyValAlaProLeuAlaGTPAsnAlaGlyLeu 92  
DB 188 CCGCACACACCG 247  
OY 93 AlaserAlaAlaInglYThrValAlaInglInglArgArgInglYglYcysAlaPheAla 112  
DB 248 GCGCGCTACCGCGACAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301  
OY 113 AspValGlyAlaSerProTyrGlyAlaAsngInglYTPAlaSerTyrArgAla----- 130  
DB 302 CACTCCCGC---GAGCCCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 358  
OY 131 ArgProAlaInglValAlaLeuTyrValAlaInglInglArgTyrTyrThrisAlaAsn 150  
DB 359 TCGGCGTCCGACCG 418  
OY 151 AsnThrcysAlaAlaGlyArgInglYcysGlyThrTyrThrgInlValAlaTPArgAsnThr 170  
DB 419 AACAGCTGCG 478  
OY 171 AlaGluValGlycysAlaInglAlaSerCysAlaThrgYala---ThleuThrLeuGly 189  
DB 479 ACCGCCATCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538  
OY 190 LeuTyrAsnProHISglYasnValGInglYglInSerProTyr 203  
DB 539 AGCTACACCG 580

RESULT 2  
US-08-181-271A-10  
Sequence 10. Application US/08181271A  
Patent No. 5614395

GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesing, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Umes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
NUMBER OF INVENTIONS: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,271A  
FILING DATE: 13-JAN-94  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-181-271A-10

Alignment Scores:  
Pred. No.: 3 12e-17 Length: 771  
Score: 304.00 Matches: 64  
Percent Similarity: 56.55% Conservative: 18  
Best Local Similarity: 44.14% Mismatches: 57  
Query Match: 28.02% Indels: 6  
DB: 1 Gaps: 3

```

US-09-832-320-2 (1-203) x US-08-181-271A-10 (1-771)
QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAla 80
Db 111 TCGCATGCCCAAACTCTCAACAGACTATTGTGGATGCCCATTAACACAGCTGTGCAGAT 170
QY 81 ValGluValAlaProLeuArgTyrAsnAlaGlyLeuAlaSerAlaAlaGlyThrVal 100
Db 171 GTAGGGGTGGAACCATTAATCTTGGGACAACGGGGTAGAGCCTATAGCAGAAATTAATGTT 230
QY 101 AlaGlnGlnArgArgGlnGlyGlyCysAlaIlePheAlaAspValGlyAlaSerProTyrGly 120
Db 231 TCGCATTTGGCTGCGAGACTGCAACCTCGTACATTCTCAT---GGCCAAATACGGC 281
QY 121 AlaAsn-----GlnGlyTyrAlaSerTyrArgAlaArgProAlaGluValAlaLeu 138
Db 282 GAAAACACGTAGCTCAGAGGAAGTGGCGAATTTATATAGC---GCTGTAAAGGGCTGCAAGATG 338
QY 139 TrpValAlaGlnGlyArgTyrTyrThrHisAlaAsnThrCysAlaAlaGlyArgGln 158
Db 339 TGGGTGCGATGAACAACACTATAGACCACTGACCAATCTTGCCACAAAGACAGGTG 398
QY 159 CysGlyThrTyrThrGlnValValTyrArgAsnThrAlaGluValAlaGlyCysAlaGlnAla 178
Db 399 TGTGGACACTATTAACACAGGTGGTGTGGCGTAACCTCGTTCGTGGATGTGCTAAGGTT 458
QY 179 SerCysAlaThrGlyAlaThrIleThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198
Db 459 AAGTGCACAACTGAGGAGATATGTTGTCTTTCGAACTATGATCTCCAGATTAATGCATA 518
QY 199 GlnGlnSerProTyr 203
Db 519 GCCCAAACTGCATAC 533

RESULT 3
US-08-449-315-10
: Sequence 10, Application US/08449315
: Patent No. 3650305
: GENERAL INFORMATION:
: APPLICANT: Ryals, John A.
: APPLICANT: Alexander, Danny C.
: APPLICANT: Beck, James J.
: APPLICANT: Duesing, John H.
: APPLICANT: Friedrich, Leslie B.
: APPLICANT: Goodman, Robert M.
: APPLICANT: Harms, Christian
: APPLICANT: Meins, Jr., Frederick
: APPLICANT: Montoya, Alice
: APPLICANT: Moyer, Mary B.
: APPLICANT: Neuhaus, Jean-Marc
: APPLICANT: Payne, George B.
: APPLICANT: Sperison, Christoph
: APPLICANT: Stinson, Jeffrey R.
: APPLICANT: Uknes, Scott J.
: APPLICANT: Ward, Eric R.
: TITLE OF INVENTION: Williams, Sherlica C.
: TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
: NUMBER OF SEQUENCES: DNA SEQUENCES AND USES THEREOF
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: CIBA-GEIGY Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: New York
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/449,315

```

```

1      FILING DATE: 24-MAY-1995
2      CLASSIFICATION: 800
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: 08/181,271
5      FILING DATE: 13-JAN-94
6      APPLICATION NUMBER: US 08/093,301
7      FILING DATE: 16-JUL-1993
8      PRIOR APPLICATION DATA:
9      APPLICATION NUMBER: US 07/937,197
10     FILING DATE: 6-NOV-1992
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: US 07/678,378
13     FILING DATE: 1-APR-1991
14     PRIOR APPLICATION DATA:
15     APPLICATION NUMBER: US 07/305,566
16     FILING DATE: 6-FEB-1989
17     PRIOR APPLICATION DATA:
18     APPLICATION NUMBER: US 07/165,667
19     FILING DATE: 8-MAR-1988
20     PRIOR APPLICATION DATA:
21     APPLICATION NUMBER: US 08/042,847
22     FILING DATE: 6-APR-1993
23     PRIOR APPLICATION DATA:
24     APPLICATION NUMBER: US 07/632,441
25     FILING DATE: 21-DEC-1990
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: US 07/425,504
28     FILING DATE: 20-OCT-1989
29     PRIOR APPLICATION DATA:
30     APPLICATION NUMBER: US 07/848,506
31     FILING DATE: 6-MAR-1992
32     PRIOR APPLICATION DATA:
33     APPLICATION NUMBER: US 07/768,122
34     FILING DATE: 27-SEP-1991
35     PRIOR APPLICATION DATA:
36     APPLICATION NUMBER: US 07/580,431
37     FILING DATE: 7-SEP-1990
38     PRIOR APPLICATION DATA:
39     APPLICATION NUMBER: US 07/368,672
40     FILING DATE: 20-JUN-1989
41     PRIOR APPLICATION DATA:
42     APPLICATION NUMBER: US 07/329,018
43     FILING DATE: 24-MAR-1989
44     PRIOR APPLICATION DATA:
45     APPLICATION NUMBER: US 08/045,957
46     FILING DATE: 12-APR-1993
47     ATTORNEY/AGENT INFORMATION:
48     NAME: Elmer, James Scott
49     REGISTRATION NUMBER: 36,129
50     REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
51     TELECOMMUNICATION INFORMATION:
52     TELEPHONE: (919)541-8614
53     TELEFAX: (919)541-8689
54     INFORMATION FOR SEQ ID NO: 10:
55     SEQUENCE CHARACTERISTICS:
56     LENGTH: 771 base pairs
57     TYPE: nucleic acid
58     STRANDEDNESS: single
59     TOPOLOGY: linear
60     MOLECULE TYPE: cDNA
61     US-08-449-315-10
62
63 Alignment Scores:
64 Pred. No.: 3,12e-17 Length: 771
65 Score: 304.00 Matches: 64
66 Percent Similarity: 56.55% Conservative: 18
67 Best Local Similarity: 44.14% Mismatches: 57
68 Query Match: 28.02% Indels: 6
69 DB: 1 Gaps: 3
70
71 US-09-832-320-2 (1-203) x US-08-449-315-10 (1-771)
72
73 61 serc15sernsnlaatrrlaaspgturyrleualapronhsaenglnalaatgalaala 800

```

DB 111 TCTCATGCCCAAACTCTCAACAAAGCTATTGTGATGCCATTAACACAGCTGTGCAGAT 170  
QY 81 ValGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSerAlaAlaGlyThrVal 100  
DB 171 GTAGCGGTGGAAACCTTACTTGGACAAAGGGGTAGACGCTATGCAAAATATGCTT 230  
QY 101 AlaGlnAlaArgAlaGlnGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
DB 231 TCTCAATGGCTGCAGACTGCACAACTCTGACATCTCAT-----GGCCAAATACGCC 281  
QY 121 AlaAsn-----GlnGlyTrpAlaSerTyrAlaArgProAlaGlyValAlaLeu 138  
DB 282 GAAACCTAGCTAGGAGAGTGGCCATTTTTANGAG---GGCTGTAAGCGCGTGCAGATG 338  
QY 139 TrpValAlaGlnGlyArgTyrTrpHisAlaAsnThrCysAlaAlaGlyArgGln 158  
DB 339 TGGGTGATGAGAAACAGTATGATGCTGCTGCACTAAATCTTGTGCACAAAGCAGGTG 398  
QY 159 CysGlyThrTyrThgAlaValAlaTrpArgAsnThrAlaGlyValGlyCysAlaGln 178  
DB 399 TGTGGACACTAATCTGAGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 458  
QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
DB 459 AAGTGCACAAATGAGGATATGTGTCTGCTGCACTATGATCTCAGGTAATGTGATA 518  
QY 199 GlyGlnSerProTyr 203  
DB 519 GGCCAAAGTCCATAC 533  
RESULT 4  
US-08-444-803-10  
; Sequence 10, Application US/08444803  
; Patent No. 5654414  
; GENERAL INFORMATION:  
; APPLICANT: Ryals, John A.  
; APPLICANT: Alexander, Danny C.  
; APPLICANT: Beck, James J.  
; APPLICANT: Duesing, John H.  
; APPLICANT: Friedlich, Leslie B.  
; APPLICANT: Goodman, Robert M.  
; APPLICANT: Harms, Christian  
; APPLICANT: Hains, Jr., Frederick  
; APPLICANT: Montoya, Alice  
; APPLICANT: Moyer, Mary B.  
; APPLICANT: Neuhaus, Jean-Marc  
; APPLICANT: Payne, George B.  
; APPLICANT: Sperison, Christoph  
; APPLICANT: Stinson, Jeffrey R.  
; APPLICANT: Umes, Scott J.  
; APPLICANT: Ward, Eric R.  
; APPLICANT: Williams, Shericoa C.  
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,803  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-444-803-10  
Alignment Scores:  
Pred. No.: 3,12e-17 Length: 771  
Score: 304.00 Matches: 64  
Percent Similarity: 56.55% Conservative: 18  
Best Local Similarity: 44.14% Mismatches: 57  
Query Match: 28.02% Indels: 6  
Gaps: 3  
US-09-832-320-2 (1-203) x US-08-444-803-10 (1-771)  
QY 61 SerGlySerAsnAlaThrAlaAspGlyTyrLeuAlaProHisAsnGlnAlaArgAla 80  
DB 111 TCTCATGCCCAAACTCTCAACAAAGCTATTGTGATGCCATTAACACAGCTGTGCAGAT 170

QY	81	ValGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal	100
Db	171	GTAGGCGGTGAAACCACTTTACTTGGGACACGCGGTACACGCTTATGCGACAAATTTATGTT	230
QY	101	AlaGlnGlnAlaGlyArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly	120
Db	231	TTCTCAATTGGCGTGCAGACTGCACACTCGTCACTTCAT-----GGCCAAATACGG	281
QY	121	AlaAsn-----GlnGlyTrpAlaSerTyrArgAlaArgProAlaGluValAlaLeu	138
Db	282	GAAAACCTAGCTCAGGGAGTGGCGCATTTATGACG---GCTCTAAGGCGCTCGAGATG	338
QY	139	TrpValAlaGlnGlyArgTyrGlyTyrTrpHisAlaAsnAsnThrCysAlaAlaGlyArgGln	158
Db	339	TGGGTCGATGAGAAACAGTACTATGACCAGTACGCAATACTCTGTGCACAGACAGCGTG	398
QY	159	CysGlyThrTyrThrGlnValValTTrpArgAsnThrAlaGluValGlyCysAlaGlnAla	178
Db	399	TGTGGACATTTACTCAGGTGGTGTGGCGTAACTCGGTTGCTGTGATGTCTAGAGTT	458
QY	179	SerCysAlaThrGlnAlaThrLeuThrLeuCysLeuTyrAsnProHisGlnAspValGln	198
Db	459	AAGTGCACATGGAGGATATGTTGTCTGTCTGCACATATGATCTCCAGGTATGTATCA	518
QY	199	GlyGlnSerProTyr	203
Db	519	GGCCAAAGTCCATAC	533
RESULT 5			
US-08-449-043-10			
:	Sequence 10, Application US/08449043		
:	Patent No. 5689044		
:	GENERAL INFORMATION:		
:	APPLICANT: Ryals, John A.		
:	APPLICANT: Alexander, Danny C.		
:	APPLICANT: Beck, James J.		
:	APPLICANT: Duesing, John H.		
:	APPLICANT: Friedrich, Leslie B.		
:	APPLICANT: Goodman, Robert M.		
:	APPLICANT: Harms, Christian		
:	APPLICANT: Meins, Jr., Frederick		
:	APPLICANT: Montoya, Alice		
:	APPLICANT: Moyer, Mary B.		
:	APPLICANT: Neuhaus, Jean-Marc		
:	APPLICANT: Payne, George B.		
:	APPLICANT: Sperison, Christoph		
:	APPLICANT: Stinson, Jeffrey R.		
:	APPLICANT: Uknes, Scott J.		
:	APPLICANT: Ward, Eric R.		
:	APPLICANT: Williams, Sherlica C.		
:	TITLE OF INVENTION: CHEMICAL REGULATABLE AND ANTI-PATHOGENIC		
:	TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF		
:	NUMBER OF SEQUENCES: 106		
:	CORRESPONDENCE ADDRESS:		
:	ADDRESSEE: CIBA-GEIGY Corporation		
:	STREET: 7 Skyline Drive		
:	City: Hawthorne		
:	STATE: New York		
:	COUNTRY: USA		
:	ZIP: 10532		
:	COMPUTER READABLE FORM:		
:	MEDIUM TYPE: Floppy disk		
:	COMPUTER: IBM PC compatible		
:	OPERATING SYSTEM: PC-DOS/MS-DOS		
:	SOFTWARE: Patentin Release #1.0, Version #1.25		
:	CURRENT APPLICATION DATA:		
:	APPLICATION NUMBER: US/08/449,043		
:	CLASSIFICATION: 800		
:	PRIOR APPLICATION DATA:		
:	APPLICATION NUMBER: 08/181,271		
:	FILING DATE: 13-JAN-94		
:	APPLICATION NUMBER: US 08/093,301		

QY 101 Alaglnclnlarqarglnlglcysalaphnealaspvalcylalaserytyrly 120  
 Db 231 TCTCAATTGGCTGCAGACGCAACCTCGTACATTCTCAT-----GGCCAAATACGGC 281  
 QY 121 Alaasn-----GlnGlyTTPAlasertyrArqAlarqProAlagluValAlaLaleu 138  
 Db 282 GAAACCTAGCTACAGGAAGTGGCGATTATATACG---CGTCTAAGGCCGCTCGAGATG 338  
 QY 139 TTPValAlagluGlyArqTyrThrHisAlaasnThrCysAlaAlaGlyArqGln 158  
 Db 339 TGGGTGATGAGAAACAGTACTATGACACAGTCAAAATCTGTGCACAGACAGGTG 398  
 QY 159 CysGlyThrTyrThrGlnValAlaTPArqAsnThrAlagluValGlyCysAlaGlnAla 178  
 Db 399 TGTGGACACTATACCTAGGTGGTGGCTTAACCTGGTTCGTGGATGTCTAGGGTT 458  
 QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
 Db 459 AAGTCAACATGAGAGATATGTCTCTTGCACACTATGATCTCCAGGTATATGCATA 518  
 QY 199 GlyGlnSerProTyr 203  
 Db 519 GGCCAAGTCCATAC 533

RESULT 6  
 US-08-456-265A-10  
 Sequence 10, Application US/08456265A  
 Patent No. 5767369

## GENERAL INFORMATION:

APPLICANT: Alexander, Danny C.  
 APPLICANT: Ryals, John A.  
 APPLICANT: Goodman, Robert M.  
 APPLICANT: Stinson, Jeffrey R.  
 TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
 TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
 NUMBER OF SEQUENCES: 111  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CIBA-GEIGY Corporation  
 STREET: 520 White Plains Road, P.O. Box 2005  
 CITY: Tarrytown  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10591

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,265A  
 FILING DATE: 31-MAY-95  
 CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/181,271  
 FILING DATE: 13-JAN-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/093,301  
 FILING DATE: 16-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/937,197  
 FILING DATE: 6-NOV-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/678,378  
 FILING DATE: 1-APR-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/305,566  
 FILING DATE: 6-FEB-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/165,667  
 FILING DATE: 8-MAR-1988  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/042,847

FILING DATE: 6-APR-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/632,441  
 FILING DATE: 21-DEC-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/425,504  
 FILING DATE: 20-OCT-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/848,506  
 FILING DATE: 6-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/768,122  
 FILING DATE: 27-SEP-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/580,431  
 FILING DATE: 7-SEP-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/368,672  
 FILING DATE: 20-JUN-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/329,018  
 FILING DATE: 24-MAR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/045,957  
 FILING DATE: 12-APR-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727/DIV10  
 TELEPHONE: (919)541-8587  
 TELEFAX: (919)541-8689  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 771 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-456-265A-10

Alignment Scores:  
 Pred. No.: 3,12e-17 Length: 771  
 Score: 304.00 Matches: 64  
 Percent Similarity: 56.55% Conservative: 18  
 Best Local Similarity: 44.14% Mismatches: 57  
 Query Match: 28.02% Indels: 6  
 Gaps: 3  
 Db: 1

US-09-832-320-2 (1-203) x US-08-456-265A-10 (1-771)

QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArqAlaAla 80  
 Db 111 TCTCATGCCCAAACTCTAACAAGACTATTTGGATGCCCATTAACAAGCTCTGTCAGAT 170  
 QY 81 ValGlyValAlaProLeuArqTTPArqAsnThrAlaGlyLeuAlaAlaGlyThrVal 100  
 Db 171 GTAGCGGTGAACATTAATTAATGCGACAGCGGGGTAGCAGCTATGACCAAAATATGTT 230  
 QY 101 Alaglnclnlarqarglnlglcysalaphnealaspvalcylalaserytyrly 120  
 Db 231 TCTCAATTGGCTGCAGACGCAACCTCGTACATTCTCAT-----GGCCAAATACGGC 281  
 QY 121 Alaasn-----GlnGlyTTPAlasertyrArqAlarqProAlagluValAlaLaleu 138  
 Db 282 GAAACCTAGCTACAGGAAGTGGCGATTATATACG---CGTCTAAGGCCGCTCGAGATG 338  
 QY 139 TTPValAlagluGlyArqTyrThrHisAlaasnThrCysAlaAlaGlyArqGln 158  
 Db 339 TGGGTGATGAGAAACAGTACTATGACACAGTCAAAATCTGTGCACAGACAGGTG 398  
 QY 159 CysGlyThrTyrThrGlnValAlaTPArqAsnThrAlagluValGlyCysAlaGlnAla 178  
 Db 399 TGTGGACACTATACCTAGGTGGTGGCTTAACCTGGTTCGTGGATGTCTAGGGTT 458



DB 399 TGTGACACTACTACAGTGGTTGGCGTACTGCGTGTGGATGCTAGGGTT 458  
 QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
 DB 459 AAGTGCACAACTAGGAGATGATGTTGCTCTTGCACAACTATGATGCTCAGATATGTCATA 518  
 QY 199 GlyGlnSerProTyr 203  
 DB 519 GCCCAAGTCCATAC 533

RESULT 7  
 US-08-455-416-10  
 : Sequence 10, Application US/08455416  
 : Patent No. 5777200  
 : GENERAL INFORMATION:  
 : APPLICANT: Ryals, John A.  
 : APPLICANT: Alexander, Danny C.  
 : APPLICANT: Beck, James J.  
 : APPLICANT: Duesing, John H.  
 : APPLICANT: Friedrich, Leslie B.  
 : APPLICANT: Goodman, Robert M.  
 : APPLICANT: Harms, Christian  
 : APPLICANT: Meins, Jr., Frederick  
 : APPLICANT: Montoya, Alice  
 : APPLICANT: Moyer, Mary B.  
 : APPLICANT: Neuhaus, Jean-Marc  
 : APPLICANT: Payne, George B.  
 : APPLICANT: Spertson, Christoph  
 : APPLICANT: Stinson, Jeffrey R.  
 : APPLICANT: Uknes, Scott J.  
 : APPLICANT: Ward, Eric R.  
 : APPLICANT: Williams, Shericca C.  
 : TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
 : TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
 : NUMBER OF SEQUENCES: 106  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: CIBA-GEIGY Corporation  
 : STREET: 7 Skyline Drive  
 : CITY: Hawthorne  
 : STATE: New York  
 : COUNTRY: USA  
 : ZIP: 10532  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/455,416  
 : FILING DATE: 31-MAY-1995  
 : CLASSIFICATION: 800  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 08/181,271  
 : FILING DATE: 13-JAN-94  
 : APPLICATION NUMBER: US 08/093,301  
 : FILING DATE: 16-JUL-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/937,197  
 : FILING DATE: 6-NOV-1992  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/678,378  
 : FILING DATE: 1-APR-1991  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/305,566  
 : FILING DATE: 6-FEB-1989  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/165,667  
 : FILING DATE: 8-MAR-1988  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/042,847  
 : FILING DATE: 6-APR-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/632,441

QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
 DB 399 TGTGACACTACTACAGTGGTTGGCGTACTGCGTGTGGATGCTAGGGTT 458  
 QY 199 GlyGlnSerProTyr 203  
 DB 519 GCCCAAGTCCATAC 533

FILING DATE: 21-DEC-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/425,504  
 FILING DATE: 20-OCT-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/848,506  
 FILING DATE: 6-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/768,122  
 FILING DATE: 27-SEP-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/580,431  
 FILING DATE: 7-SEP-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/368,672  
 FILING DATE: 20-JUN-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/329,018  
 FILING DATE: 24-MAR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/045,957  
 FILING DATE: 12-APR-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919)541-8614  
 TELEFAX: (919)541-8689  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 771 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 US-08-455-416-10

Alignment Scores:  
 Pred. No.: 3,12e-17 Length: 771  
 Score: 304.00 Matches: 64  
 Percent Similarity: 56.55% Conservative: 18  
 Best Local Similarity: 44.14% Mismatches: 57  
 Query Match: 28.02% Indels: 6  
 DB: 1 Gaps: 3

US-09-832-320-2 (1-203) x US-08-455-416-10 (1-771)  
 QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAla 80  
 DB 111 TCTCATGGCCAAACCTCTCACAACAGACTATTGATGCCCATACACAGCTCGCAGAT 170  
 QY 81 ValGlyValAlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaAlaGlyThrVal 100  
 DB 171 GTAGCGGTGGAACCTTAACTTGGGACACAGCGGTACACGCTTGCACAAATTATGTT 230  
 QY 101 AlaGlnAlaArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
 DB 231 TCTCAATTGGCTGCGAGACTGCAACCTGCTACATTCTCAT-----GCCCATACGCC 261  
 QY 121 AlaAsn-----GlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaLeu 138  
 DB 282 GAAACCTTAGCTCAGGAGAGTGGCGATTATTATGACG--GCTGCTAGCGCTCGAGATG 338  
 QY 139 TrpValAlaGlnGlyArgTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGln 158  
 DB 339 TGGGTGATGAGAACACTATGACCATATGACCAATACTTGTGCACAGACAGGTG 398  
 QY 159 CysGlyThrTyrThrGlnValValTyrPargAsnThrAlaGluValGlyCysAlaGlnAla 178  
 DB 399 TGTGACACTATACACAGTGGTTGGCGTAACTTCGCTGCTGTTGATGCTAGGGTT 458  
 QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198

DB 459 AAGTGAACAAATGAGATATGTCTCTGCAACTATGATCTCCAGGTAATGTCTATA 518  
OY 199 GYGLInserProTyr 203  
DB 519 GGCCAAAGTCATAC 533

RESULT 8  
US-08-455-244-10  
; Sequence 10, Application US/08455244  
; Patent No. 5789214

GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Alexander, Danny C.  
APPLICANT: Beck, James J.  
APPLICANT: Duesling, John H.  
APPLICANT: Friedrich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Sheriecca C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF SEQUENCES: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,244  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-94  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504

FILING DATE: 20-OCT 1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018  
FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/PL/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-455-244-10

Alignment Scores:  
Pred. No.: 3,12e-17 Length: 771  
Score: 304.00 Matches: 64  
Percent Similarity: 56.55% Conservative: 18  
Best Local Similarity: 44.14% Mismatches: 57  
Query Match: 28.02% Indels: 6  
DB: 1 Gaps: 3

US-09-832-320-2 (1-203) x US-08-455-244-10 (1-771)

OY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaAla 80  
DB 111 TCTCATGGCCAAACCTCTCAACAAAGCATTTGGATGCCCATACACAGCTGTCAGAT 170  
OY 81 ValGlyValAlaProLeuArgTyrAsnAlaGlyLeuAlaSerAlaAlaAlaGlyTyrVal 100  
DB 171 GTAGGGGTGGAACCATTTACTTTGGACAACGGGTGACGACCTATGCACAAATTTATGTT 230  
OY 101 AlAGlnGlnArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
DB 231 TCTCAATTGGCTGACAGACTGCAACCTCGTACATCTCAT-----GGCCAAATACGGC 281  
OY 121 AlaAsn-----GlnGlyTyrPalaSerTyrArgAlaArgProAlaGlnValAlaLeu 138  
DB 282 GAACACCTAGCTCAGGAGAGTGGCGATTTTATAGCG---GCTGTAGGCGCTGAGATG 338  
OY 139 TrpValAlaGlnGlyArgTyrTyrThrHisAlaAsnThrCysAlaAlaGlyArgGln 158  
DB 339 TGGGTGATGAGAAACAGTACTATGACCTGACCTAAATACTGTGTGCACAAAGACAGTG 398  
OY 159 CysGlyTyrTyrThrGlnValValTyrPargAsnThrAlaGlnValGlyCysAlaGlnAla 178  
DB 399 TGTGACACTATACCTAGGTGTTGGCGTACCTGCGTTGCGTTGATGTCGTAGGTT 458  
OY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
DB 459 AAGTGAACAAATGAGATATGTGTCTCTGCAACTATGATCTCCAGGTAATGTCTATA 518





```

: Sequence 10, Application US/08456262
: Patent No. 5851766
: GENERAL INFORMATION:
: APPLICANT: Kyalis, John A.
: APPLICANT: Alexander, Danny C.
: APPLICANT: Beck, James J.
: APPLICANT: Duesing, John H.
: APPLICANT: Friedrich, Leslie B.
: APPLICANT: Goodman, Robert M.
: APPLICANT: Harms, Christian
: APPLICANT: Meins, J., Frederick
: APPLICANT: Montoya, Alice
: APPLICANT: Moyer, Mary B.
: APPLICANT: Neuhaus, Jean-Marc
: APPLICANT: Payne, George B.
: APPLICANT: Sperison, Christoph
: APPLICANT: Stinson, Jeffrey R.
: APPLICANT: Uknes, Scott J.
: APPLICANT: Ward, Eric R.
: APPLICANT: Williams, Shericea C.
: TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
: TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CIBA-GEIGY Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: New York
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/456,262
: FILING DATE: 31-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/181,271
: FILING DATE: 13-JAN-94
: APPLICATION NUMBER: US 08/093,301
: FILING DATE: 16-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/937,197
: FILING DATE: 6-NOV-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/678,378
: FILING DATE: 1-APR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/305,566
: FILING DATE: 6-FEB-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/165,667
: FILING DATE: 8-MAR-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/042,847
: FILING DATE: 6-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/632,441
: FILING DATE: 21-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/425,504
: FILING DATE: 20-OCT-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/848,506
: FILING DATE: 6-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/768,122
: FILING DATE: 27-SEP-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/580,431

```

```

: FILING DATE: 7-SEP-1990
: APPLICATION DATA:
: APPLICATION NUMBER: US 07/368,672
: FILING DATE: 20-JUN-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/329,018
: FILING DATE: 24-MAR-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/045,957
: FILING DATE: 12-APR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919)541-8614
: TELEFAX: (919)541-8689
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-456-262-10

Alignment Scores:
Pred. No.: 3 128-17 Length: 771
Score: 304.00 Matches: 64
Percent Similarity: 56.35% Conservative: 18
Best Local Similarity: 44.14% Mismatches: 57
Query Match: 28.02% Indels: 6
DB: Gaps: 3

US-09-832-320-2 (1-203) x US-08-456-262-10 (1-771)
QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaAla 80
DB 111 TCTCATGCCCCAAACCTCTCAACAGACTATTTGGATGGCCATTAACACAGCTCGCAGAT 170
QY 81 ValGlyValAlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal 100
DB 171 GTAGGCGTGAACCATTAAGTCTGGGCAACGGGGTAGCAGCCATGACAAATTTATGTT 230
QY 101 AlaGlnGlnArgArgGlnGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120
DB 231 TCTCAATTGGCTGCTGACAGCTGCAACTCTGATCTTCAT-----GGCCAAATAGGC 281
QY 121 AlaAsn-----GlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaAlaLeu 138
DB 282 GAACACCTAGCTCAGGAGAGTGGCGATTATTATGACG--GCTGCTAAGGCCGTCGAGATG 338
QY 139 TyrValAlaGlnGlyArgTyrTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGln 158
DB 339 TGGGTGATAGAAACAGAGTCTGATGACATGACATGATATCTTGTGCAACAGACAGAGTG 398
QY 159 CysGlyThrTyrThrGlnValAlaTyrPargAsnThrAlaGluValGlyCysAlaGlnAla 178
DB 399 TGTGACACATATCTACAGGTGGTGGCTTACCTGCTGTGGATGTGCTAGAGCTT 458
QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198
DB 459 AAGTGCACAAATGAGGAGATATGTGTCTCTTGCACACTATGATCTCCAGGATATGTCATA 518
QY 199 GlyGlnSerProTyr 203
DB 519 GGCCTAAGTCCATAG 533

RESULT 12
US-08-456-240-10
: Sequence 10, Application US/08456240
: Patent No. 585154
: GENERAL INFORMATION:

```

```

      FILING DATE: 20-JUN-1989
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 07/329,018
        FILING DATE: 24-MAR-1989
        PRIOR APPLICATION DATA:
          APPLICATION NUMBER: US 08/045,957
          FILING DATE: 12-APR-1993
          ATTORNEY/AGENT INFORMATION:
            NAME: Elmer, James Scott
            REGISTRATION NUMBER: 36,129
            REFERENCE/DOCKET NUMBER: S-19825/PJ/CSC 1727
            TELECOMMUNICATION INFORMATION:
              TELEPHONE: (919)541-8614
              TELEFAX: (919)541-8689
            INFORMATION FOR SEQ ID NO: 10:
              SEQUENCE CHARACTERISTICS:
                LENGTH: 771 base pairs
                TYPE: nucleic acid
                STRANDEDNESS: single
                TOPOLOGY: linear
              MOLECULE TYPE: cDNA
      US-08-456-240-10

Alignment Scores:
Pred. No.:           3,12e-17             Length:       771
Score:             304.00               Matches:      64
Percent Similarity: 56.55%              Conservative: 18
Best Local Similarity: 44.14%            Mismatches:   57
Query Match:       26.02%               Indels:       6
                                      Gaps:         3

US-09-832-320-2 (1-203) x US-08-456-240-10 (1-771)

QY      61  SerGISeRnaAlaThrAlaAspLyuTrLeuAlaProHIsAsnGlnAlaArgAlaIa  80
      |||:::|||||:::|||||  |||||  |||||
Db      111 TCTCATGCCCAAAACGTCTAACAGACATTTCGATGCCCAATACAGACTCTGCAGT  178

QY      81  ValGIyValaIaProlEuArTrIPhsnRIgLYleuAlaSerAlaIaIaIgLythrVal  100
      |||||:::|||||  |||:::|||||:::|||||  |||
Db      171 GRAGCGGTGGAGAACCTTAATCTTGGSACACAGGGGTGACGCTTAGCATAAATTATGT  230

QY      101 AlaGInGlnArGrArGlInGLyGLcysAlaPheAlaSerValGIyAlaSerProtyrGLy  120
      ::|||  :::  :::
Db      231 TCTCAATTGGCTGCAGACACGTCAACCTCGTACATTTCCAT-----GGCCAAATPAGGC  281

QY      121 AlaaAn-----GInGLyTrPaIaSerTyArGIalArGProlaIGluValaIaLeu  138
      |||  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      282 GAAGAACCtAGtCGAcGGAAGtGGCGAtTTTATAGC---GCTGCTAAGCCGCTCGAGATG  338

QY      139 TrPVaIaIaGluGLyArqTYrTYrThNHisAlaAsnThrCYsaIaIaGlyArgLn  156
      |||||  |||  :::|||||  |||  :::|||||  |||  |||  |||  |||  |||  |||
Db      339 TGCGTCGATGGAAACAGTACATACCACTCAAGTCAAAATACCTTGTCGACMAAGCAGGTG  398

QY      159 CySGlYThrTYrThrGlnValaITrPArASnTHrAlaIGluValaIGlyCYsaIaGlnAla  178
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      399 TGTGACACtATACtACtAGAGTGTGTTGGGTGATACtCGTTCGHTGGATGTCTAGAGTT  458

QY      179 SerCYsaIaThrTrGlyAlaThrLeuThrLeuCysLeuTYrAAsnProHISglYAsnValGln  196
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      459 AAGTCAACAATGGAGAGATATGTCTCTTGCAACTATGATCTCCAGSTATTATGATA  518

QY      199 gLyGInserProtyr  203
      |||||  |||||  |||||
Db      519 GCCCAAGTCCATAC  533

RESULT 13
US-08-455-736-10
Sequence 10, Application US/08455736
Patent No. 5880328
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
```

APPLICANT: Duesing, John H.  
APPLICANT: Friedlich, Leslie B.  
APPLICANT: Goodman, Robert M.  
APPLICANT: Harms, Christian  
APPLICANT: Meins, Jr., Frederick  
APPLICANT: Montoya, Alice  
APPLICANT: Moyer, Mary B.  
APPLICANT: Neuhaus, Jean-Marc  
APPLICANT: Payne, George B.  
APPLICANT: Sperison, Christoph  
APPLICANT: Stinson, Jeffrey R.  
APPLICANT: Uknes, Scott J.  
APPLICANT: Ward, Eric R.  
APPLICANT: Williams, Shericea C.  
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC  
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/455,736  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/181,271  
FILING DATE: 13-JAN-1994  
APPLICATION NUMBER: US 08/093,301  
FILING DATE: 16-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/937,197  
FILING DATE: 6-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,378  
FILING DATE: 1-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/305,566  
FILING DATE: 6-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/165,667  
FILING DATE: 8-MAR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,847  
FILING DATE: 6-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/632,441  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/425,504  
FILING DATE: 20-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/848,506  
FILING DATE: 6-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/768,122  
FILING DATE: 27-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/580,431  
FILING DATE: 7-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/368,672  
FILING DATE: 20-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/329,018

FILING DATE: 24-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,957  
FILING DATE: 12-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-455-736-10  
Alignment Scores:  
Pred. No.: 3,12e-17 Length: 771  
Score: 304.00 Matches: 64  
Percent Similarity: 56.55% Conservative: 18  
Best Local Similarity: 44.14% Mismatches: 57  
Query Match: 28.02% Indels: 6  
Gaps: 3  
US-09-832-320-2 (1-203) x US-08-455-736-10 (1-771)  
QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaAla 80  
Db 111 TCTCATGCCCAAACTCTCAACAGACTATTGGATGCCCAACAGCTCGTCACAT 170  
QY 81 ValGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSerAlaAlaGlyThrVal 100  
Db 171 GTAGCGGTGACACCTATTACCTGGACACAGGGGTGACGACCTATGACCAAAATTATGTT 230  
QY 101 AlaGlnAlaArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
Db 231 TCTCAATTGGCTGCAGACTCAACCTGTAACATTCAT-----GCCCAATACGGC 281  
QY 121 AlaAsn-----GlnGlyTrpAlaSerTyrArgAlaArgProAlaGlnValAlaAla 138  
Db 282 GAATACCTAGCTCAGGAGAGTGGCGATTATGACG---GCTGCTAAGCGCTCGACATG 338  
QY 139 TrpValAlaGlnGlyArgTyrTrpHisAlaAsnAsnThrCysAlaAlaGlyArgGln 158  
Db 339 TCGGTGCATGAGAAACAGTACTATGACATGACTCAATATCTGTGCACAGACAGGTG 398  
QY 159 CysGlyThrTyrTrpGlnValAlaTrpArgAsnThrAlaGlnValGlyCysAlaGlnAla 178  
Db 399 TGTGACACTATCTCAGGTGGTGTGGCTGACTCGCTTCGTGGATGTCTAGGGTT 458  
QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
Db 459 AAGTCAACAATGAGAGATATGTTGCTCTTGCACACTATGATCCCTCCAGTATATGCATA 518  
QY 199 GlyGlnSerProTyr 203  
Db 519 GGCCAAAGTCCATAC 533  
RESULT 14  
US-08-971-217-10  
Sequence 10, Application US/08971217  
Patent No. 5942662  
GENERAL INFORMATION:  
APPLICANT: Ryals, John A.  
APPLICANT: Harms, Christian  
APPLICANT: Friedlich, Leslie  
APPLICANT: Beck, James  
APPLICANT: Uknes, Scott  
APPLICANT: Ward, Eric









GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 22, 2003, 03:04:05 : Search time 106 Seconds  
(without alignments)  
1075.592 Million cell updates/sec

Title: US-09-832-320-2  
Perfect score: 1085  
Sequence: 1 MAHSRSHHLLLPAPMATA.....ATLTLCLYHNGVNGOSPY 203

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 442118 seqs, 280819700 residues  
Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USPPO1/US09832320/runat.20022003.092738.14846/app.query.fasta.1.391  
-DB=Published.Applications.NA -OPMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAPPIT=blosum62  
-TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEATSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09832320\_@CGN\_1.80\_@runat.20022003.092738.14846  
-DEVT=TIMEOUT=120 -WARP=NO\_XLPPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications.NA.\*  
1: /cgn2\_6/pdata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/pdata/2/pubpna/PCT\_NEM\_PUB.seq:\*  
3: /cgn2\_6/pdata/2/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/pdata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/pdata/2/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/pdata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/pdata/2/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/pdata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/pdata/2/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/pdata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/pdata/2/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cgn2\_6/pdata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/pdata/2/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cgn2\_6/pdata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1085	100.0	612	10	US-09-832-320-3
2	1085	100.0	898	10	US-09-832-320-1
3	333	30.7	749	10	US-09-840-479-12
4	300.5	27.7	772	9	US-10-068-347-3

5	297.5	27.4	492	9	US-09-938-842A-1676	Sequence 1676, Ap
6	294	27.1	501	9	US-09-938-842A-2002	Sequence 2002, Ap
7	292.5	27.0	866	10	US-09-840-479-6	Sequence 6, Appl1
8	288	26.5	486	10	US-09-938-842A-2047	Sequence 2047, Ap
9	288	26.5	738	10	US-09-887-576-608	Sequence 608, App
10	288	26.5	757	10	US-09-887-576-618	Sequence 618, App
11	287	26.5	714	12	US-10-078-929-85	Sequence 85, Appl
12	280.5	25.9	604	12	US-10-078-929-99	Sequence 99, Appl
13	264	24.3	783	12	US-10-078-929-81	Sequence 81, Appl
14	264	24.3	806	10	US-09-840-479-14	Sequence 14, Appl
15	255	23.5	701	12	US-10-078-929-95	Sequence 95, Appl
16	254	23.4	939	12	US-10-078-929-89	Sequence 89, Appl
17	249	22.9	547	12	US-10-078-929-97	Sequence 97, Appl
18	248	22.9	718	12	US-10-078-929-93	Sequence 93, Appl
19	236	21.8	534	12	US-10-078-929-83	Sequence 83, Appl
20	231	21.3	664	9	US-10-068-347-1	Sequence 1, Appl1
21	223.5	20.6	525	9	US-10-068-347-9	Sequence 9, Appl1
22	205	18.9	1876	9	US-09-944-433-49	Sequence 49, Appl
23	205	18.9	1876	9	US-09-944-403-49	Sequence 49, Appl
24	205	18.9	1876	9	US-09-944-896-49	Sequence 49, Appl
25	205	18.9	1876	9	US-09-944-944-49	Sequence 49, Appl
26	205	18.9	1876	9	US-09-944-907-49	Sequence 49, Appl
27	205	18.9	1876	9	US-09-944-929-49	Sequence 49, Appl
28	205	18.9	1876	10	US-09-866-028-49	Sequence 49, Appl
29	205	18.9	1876	10	US-09-944-449-49	Sequence 49, Appl
30	205	18.9	1876	10	US-09-944-457-49	Sequence 49, Appl
31	205	18.9	1876	10	US-09-945-587-49	Sequence 49, Appl
32	205	18.9	1876	10	US-09-945-015-49	Sequence 49, Appl
33	205	18.9	1876	10	US-09-944-396-49	Sequence 49, Appl
34	205	18.9	1876	10	US-09-944-097-49	Sequence 49, Appl
35	205	18.9	1876	10	US-09-944-432-49	Sequence 49, Appl
36	205	18.9	1876	10	US-09-943-762-49	Sequence 49, Appl
37	205	18.9	1876	10	US-09-943-654-49	Sequence 49, Appl
38	205	18.9	1876	10	US-09-943-851A-49	Sequence 49, Appl
39	202	18.6	360	10	US-09-770-791-523	Sequence 523, App
40	200	18.4	1338	10	US-09-780-264-3	Sequence 3, Appl1
41	200	18.4	1856	10	US-09-780-264-1	Sequence 1, Appl1
42	200	18.4	1923	9	US-10-042-141-12	Sequence 12, Appl
43	200	18.4	1923	10	US-09-726-643-12	Sequence 12, Appl
44	197	18.2	483	9	US-09-938-842A-1683	Sequence 93, Appl
45	196	18.1	472	12	US-10-078-929-91	Sequence 91, Appl

#### ALIGNMENTS

RESULT 1  
US-09-832-320-3  
Sequence 3, Application US/09832320  
Patent No. US20010049834A1  
GENERAL INFORMATION:  
APPLICANT: Crane, Edmund H.  
TITLE OF INVENTION: Maltase pathogenesis-Related  
FILE REFERENCE: 35718/214291  
CURRENT APPLICATION NUMBER: US/09/832,320  
CURRENT FILING DATE: 2001-04-10  
PRIOR APPLICATION NUMBER: US 60/195,801  
PRIOR FILING DATE: 2000-04-10  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 612  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(612)  
US-09-832-320-3  
Alignment Scores:  
Pred. No.: 9,86e-88 Length: 612  
Score: 1085.00 Matches: 203

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-832-320-2 (1-203) x US-09-832-320-3 (1-612)

QY 1 MetalaHisSerArgSerHisSHSHSLeuLeuLeuProAlaProMetaIaThraIa 20  
 DB 1 ATGGGCGACTGCGCAGCCACCCACCTCTGCGTCCGCGCCGCGCCGCGCCGCGCGCGG 60  
 QY 21 CysteuleuLeuAlaThrLeuLeuAlaLeuCysAlaAlaProAlaProThrhISglYAla 40  
 DB 61 TCGTTCCTCTGCGCAGCCCTCTGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGG 120  
 QY 41 ArgValIleuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyThrGly 60  
 DB 121 CGCGTCCATGCGCGGCGCGCGCGCGCGCGGTGACCAAGCGCAGCGGTGCGCAGCGG 180  
 QY 61 SerGlySerAsnAlaThraIaAspGluTyrLeuAlaProHisAsnGlnAlaIaIa 80  
 DB 181 AGCGGCGACCAAGCGGCGCGGAGAGTACCTGCGCGCGCAACCGCGCGCGCGCGCGG 240  
 QY 81 ValGlyValAlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaIaAlaIaGlyThrVal 100  
 DB 241 GTGGGCGTGGCCCGCTGCGGTGGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGG 300  
 QY 101 AlaGlnGlnArgArgGlnGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
 DB 301 GCGGCGCGCGCGCGCGCGCGCGGTGCGCGTTCGCGGAGCGTGGGCGCGCGCGCTACGCG 360  
 QY 121 AlaAsnGlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaIaLeuTyrVal 140  
 DB 361 GCGAACCGAGGGGTGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGTGTGGG 420  
 QY 141 AlaGlnGlyArgTyrTyrThrHisAlaAsnThrCysAlaAlaGlyArgGlnCysGly 160  
 DB 421 GCGGAGGGCGGTACTACACCGCGCAACACGAGTGGCGCGCGCGCGCGCGGTGCGG 480  
 QY 161 ThrTyrThrGlnValValTyrPargAsnThraIaGluValGlyCysAlaGlnAlaSerCys 180  
 DB 481 AGGTACAGCAGGTGTGTGGCGAACCGCGCGAGGTGGGTGGCGCGCGCGCGCGTGC 540  
 QY 181 AlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGln 200  
 DB 541 GCGACGCGCGCGCGCGCTACCGCTGCGCTGTACAAACCGCGCAGCGCAGCGCGCGAG 600  
 QY 201 SerProTyr 203  
 DB 601 AGCCCTCTAC 609

RESULT 2

US-09-832-320-1

; Sequence 1, Application US/09832320

; Patent No. US20010049834A1

; GENERAL INFORMATION:

; APPLICANT: Crane, Edmund H.

; APPLICANT: Crane, Virginia C.

; TITLE OF INVENTION: Maize Pathogenesis-Related

; FILE REFERENCE: 35718/214291

; CURRENT APPLICATION NUMBER: US/09/832,320

; CURRENT FILING DATE: 2001-04-10

; PRIOR APPLICATION NUMBER: US 60/195,801

; PRIOR FILING DATE: 2000-04-10

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 898

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (63)...(674)

US-09-832-320-1

Alignment Scores:

Pred. No.: 1,56e-87

Score: 1085.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 10

Length: 898

Matches: 203

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-832-320-2 (1-203) x US-09-832-320-1 (1-898)

QY 1 MetalaHisSerArgSerHisSHSHSLeuLeuLeuProAlaProMetaIaThraIa 20  
 DB 63 ATGGGCGACTGCGCAGCCACCCACCTCTGCGTCCGCGCCGCGCGCGCGCGCGCGG 122  
 QY 21 CysteuleuLeuAlaThrLeuLeuAlaLeuCysAlaAlaProAlaProThrhISglYAla 40  
 DB 123 TCGTTCCTCTGCGCAGCCCTCTGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGG 182  
 QY 41 ArgValIleuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyThrGly 60  
 DB 183 CGCGTCCATGCGCGGCGCGCGCGCGCGGTGACCAAGCGCAGCGGTGCGCAGCGG 242  
 QY 61 SerGlySerAsnAlaThraIaAspGluTyrLeuAlaProHisAsnGlnAlaIaIa 80  
 DB 243 AGCGGCGACCAAGCGGCGCGGAGAGTACCTGCGCGCGCAACCGCGCGCGCGCGG 302  
 QY 81 ValGlyValAlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaIaAlaIaGlyThrVal 100  
 DB 303 GTGGGCGTGGCCCGCTGCGGTGGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGG 362  
 QY 101 AlaGlnGlnArgArgGlnGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
 DB 363 GCGGAGCGCGCGCGCGCGCGGTGCGCGTTCGCGGAGCGTGGGCGCGCGCGCTACGCG 422  
 QY 121 AlaAsnGlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaIaLeuTyrVal 140  
 DB 423 GCGAACCGAGGGGTGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGTGTGGG 482  
 QY 141 AlaGlnGlyArgTyrTyrThrHisAlaAsnThrCysAlaAlaGlyArgGlnCysGly 160  
 DB 483 GCGGAGGGCGGTACTACACCGCGCAACACGAGTGGCGCGCGCGCGCGCGGTGCGG 542  
 QY 161 ThrTyrThrGlnValValTyrPargAsnThraIaGluValGlyCysAlaGlnAlaSerCys 180  
 DB 543 AGGTACAGCAGGTGTGTGGCGAACCGCGCGAGGTGGGTGGCGCGCGCGCGCGTGC 602  
 QY 181 AlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGln 200  
 DB 603 GCGACGCGCGCGCGCGCTACCGCTGCGCTGTACAAACCGCGCAGCGCAGCGCGAG 662  
 QY 201 SerProTyr 203  
 DB 663 AGCCCTCTAC 671

RESULT 3

US-09-840-479-12

; Sequence 12, Application US/09840479

; Patent No. US20010025380A1

; GENERAL INFORMATION:

; APPLICANT: Crane, Virginia

; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters

; FILE REFERENCE: 5718-32, 035718/175219

; CURRENT APPLICATION NUMBER: US/09/840,479

; CURRENT FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: 09/257,583

; PRIOR FILING DATE: 1999-02-25

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 749

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (92)..(583)
US-09-840-479-12

Alignment Scores:
Pred. No.: 2,71e-21 Length: 749
Score: 333.00 Matches: 81
Percent Similarity: 48.13% Conservative: 22
Best Local Similarity: 37.85% Mismatches: 67
Query Match: 30.69% Indels: 44
DB: Gaps: 8

US-09-832-320-2 (1-203) x US-09-840-479-12 (1-749)

QY 5 ArgSerHisHisLeuLeuLeuProAla----- 15
   |||||  |||  |||  |||
Db 26 CGATCAGACATTTGGACTTGGCTGCTGCTCATTAATTACTTACTTATCGCAACAA 85

QY 16 ---Prometala-----ThralaCysLeuLeuAlaThrLeuAlaLeuCysAla 32
   |||||  |||||  |||||  |||||
Db 86 CAACAATGGCAGCGAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 145

QY 33 AlaProAlaProThrHisGlyAlaArgValLeuMetProGlyGlyAlaValThr 52
   |||||  |||||  |||||  |||||
Db 146 GCGCGGTGC----- 154

QY 53 LysAlaGlnGlnGlyGlyThrGlySerGlySerAlaThrAlaAspGluThrLeuAla 72
   |||||  |||||  |||||  |||||
Db 155 -----ACGCCCGCAAGACTGCGCGCAGAC---TACGTGAC 187

QY 73 ProHisAsnGlnAlaArgAlaAlaValAlaProLeuArgTrpAsnAlaGlyLeu 92
   |||||  |||||  |||||  |||||
Db 188 CGCAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247

QY 93 AlaSerAlaAlaAlaGlyThrValAlaGlnGlnAlaArgGlnGlnGlyCysAlaPheAla 112
   |||||  |||||  |||||  |||||
Db 248 GCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301

QY 113 AspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTrpAlaSerTyrArgAla----- 130
   |||||  |||||  |||||  |||||
Db 302 CACGTGCGGCG-----GCGCGCTAGCGCGCAAGACCTGCTGCGCGCGCGCGCGCGCG 358

QY 131 ArgProAlaGlnValAlaLeuTrpValAlaGlnGlyArgTyrTrpHisAlaAsn 150
   |||||  |||||  |||||  |||||
Db 359 TCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418

QY 151 AsnThrCysAlaAlaGlyArgGlnCysGlyThrTrpThrGlnValAlaTrpArgAsnThr 170
   |||||  |||||  |||||  |||||
Db 419 AACACAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478

QY 171 AlaGlnValGlyCysAlaGlnAlaSerCysAlaThrGlyAla---ThrLeuThrLeuCys 189
   |||||  |||||  |||||  |||||
Db 479 ACCGCGATCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538

QY 190 LeuTyrAsnProHisGlyAsnValGlnGlnSerProTyr 203
   |||||  |||||  |||||  |||||
Db 539 AGCTACAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 580

RESULT 4
US-10-068-347-3
; Sequence 3, Application US/10068347
; Patent No. US200201616A1
; GENERAL INFORMATION:
; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Acevedo, Pedro
; APPLICANT: Simmons, Carl
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Maize P1 Polynucleotides and Methods of Use
; FILE REFERENCE: 35718/242798 (5718-151)
; CURRENT APPLICATION NUMBER: US/10/068,347
```

```

; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/267,052
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(571)
US-10-068-347-3

Alignment Scores:
Pred. No.: 2,07e-18 Length: 772
Score: 300.50 Matches: 70
Percent Similarity: 52.10% Conservative: 17
Best Local Similarity: 41.92% Mismatches: 63
Query Match: 27.70% Indels: 17
DB: Gaps: 7

US-09-832-320-2 (1-203) x US-10-068-347-3 (1-772)

QY 42 ValLeuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyThrGlySer 61
   |||||  |||||  |||||  |||||
Db 104 GTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 148

QY 62 GlySerAsnAlaThrAlaAspGluThrLeuAlaProHisAsnGlnAlaAlaValAla 81
   |||||  |||||  |||||  |||||
Db 149 -----ACGCGCAGAGACTTGTGAACCTGCACAACTCCCGCGCGGAGCGTG 196

QY 82 GlyValAlaProLeuAlaArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAla 101
   |||||  |||||  |||||  |||||
Db 197 GCGCTGGGAAGCTGCGCTGGAAACCCACGCTGCGC---GCTACGCGCAGACTGCGCG 253

QY 102 GlnGlnAlaArgGlnGlnGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAla 121
   |||||  |||||  |||||  |||||
Db 254 AACGAG---GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307

QY 122 AsnGlnGlyTrp-----AlaSerTyrArgAlaArgProAlaGlnValAlaLeuTrp 139
   |||||  |||||  |||||  |||||
Db 308 AACCTGTTCTGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367

QY 140 ValAlaGlnGlyArgTyrTrpHisAlaAsnThrCysAlaAla-----GlyArg 157
   |||||  |||||  |||||  |||||
Db 368 GCGCGGAGAAGAGTACTCAACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 427

QY 158 GlnCysGlyThrTrpThrGlnValAlaTrpArgAsnThrAlaGlnValGlyCysAlaGln 177
   |||||  |||||  |||||  |||||
Db 428 TCGTGGCGCACTACAGCAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487

QY 178 AlaSerCysAlaThrGlyAla---ThrLeuThrLeuCysLeuTyrAsnProHisGlyAsn 196
   |||||  |||||  |||||  |||||
Db 488 GTGCTGCGCAGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547

QY 197 ValGlnGlnGlnSerProTyr 203
   |||||  |||||  |||||  |||||
Db 548 GTGATTTGACAGAGACCCCTTAC 568

RESULT 5
US-09-938-842A-1676
; Sequence 1676, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harter, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
```

```

; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO: 1676
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1676

```

## Alignment Scores:

```

Pred. No.: 2,21e-18 Length: 492
Score: 297.50 Matches: 61
Percent Similarity: 57.78% Conservative: 17
Best Local Similarity: 45.19% Mismatches: 34
Query Match: 27.42% Indels: 3
DB: Gaps: 2

```

US-09-832-320-2 (1-203) x US-09-938-842A-1676 (1-492)

```

QY 69 GUTYRLTAUAIARONIAASNGIALAAGALAIAlaValGlyAlaIAlaProLeuArgTrp 88
DB 94 GACTTTTGGGAGTTTCACACCGAGCGAGCGAGGTTGGGAGCCCTTAAGATGG 153
QY 89 ASNAIAGIYLeuAlaSerAlaAlaGlyThrValAlaGlnGlnArgArgGlnGly 108
DB 154 GACGAGAAAGGTGGCTGTATGCGCGTAACTAGCTAACAG-----CGTAAAGGTGAC 207
QY 109 CYSAlaIAlaPheAlaAspValAlaGlyAlaSerProtyrGlyAlaAsnGlnGlyTrpAlaSerTrp 128
DB 208 TCGGCTATGAAACACTCGACGGGGTCC--TATGAGAGAAACATCGCTTGGACACGGGT 264
QY 129 ATGAlaArgProAlaGlyValAlaAlaLeuTrpValAlaGlnGlyArgGlyTrpHis 148
DB 265 AACATGACAGCGGTGACGAGCAAGTTGACATGTGGTGGCGAGCAATTGACATGATTTAT 324
QY 149 AlAaAsnAsnThrCysAlaAlaAlaGlyArgGlnGlySerGlyThrTrpGlnValAlaTrpArg 168
DB 325 GATTCACAAATCATGTGCTATGGACAAACATGTCGCCACTATGCTGGTGTGGAGA 384
QY 169 ASnThrAlaGlyValAlaGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThrLeu 188
DB 385 AACTCGAGAGGTGGTGGATGTGCAAAAGTGAATGCAACAATGGTCAACTTATACACT 444
QY 189 CYSLeuTrpAsnProHisGlyAsnValGlnGlyInsProtyr 203
DB 445 TGCACACTACGATCCCGGCTAACACTGGGTGGCGAGTGGCCCTTAC 489

```

RESULT 6  
US-09-938-842A-2002  
; Sequence 2002, Application US/09938842A  
; Patent No. US20020160378A1

```

; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

```

```

; SEQ ID NO 2002
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-2002

```

## Alignment Scores:

```

Pred. No.: 4.61e-18 Length: 501
Score: 294.00 Matches: 70
Percent Similarity: 47.09% Conservative: 19
Best Local Similarity: 37.04% Mismatches: 72
Query Match: 27.10% Indels: 28
DB: Gaps: 4

```

US-09-832-320-2 (1-203) x US-09-938-842A-2002 (1-501)

```

QY 15 AlAProMetAlaThrAlaCysLeuLeuAlaThrLeuLeuAlaLeuCysAlaAlaPro 34
DB 16 TCTCTCAAAACCTTACTACTATGCGATTAGCTTGTCTCACTTTT--GCAAGTTCA 72
QY 35 AlAProThrHisGlyAlaArgValLeuMetProGlyAlaGlyAlaValThrLysAla 54
DB 73 -----CTGAAAGCC 81
QY 55 GlnGlnGlyGlyThrGlySerGlySerAsnAlaThrAlaAspGlyTrpLeuAlaProHis 74
DB 82 CAAAGAC-----CGCGCAAGAACTACTCTGATGTACAC 114
QY 75 ASnGlnAlaArgAlaAlaValAlaGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSer 94
DB 115 AACCATGCTGAGAGCGAGCTTACTAGTCTCATATTAATATGATGCGGGA--GCAAGCC 171
QY 95 AlAAlaAlaGlyThrValAlaGlnGlnArgArgGlnGlyCysAlaIAlaPheAlaAspVal 114
DB 172 CGATACCGCTGGAACATATGCTCAAGAGAAAGCGGAGCTGTGCTCATTCACCTCAAC 231
QY 115 GlyAlaSerProtyrGlyAlaAsnGlnGlyTrpAlaSerTrpArgAlaArgProAlaGly 134
DB 232 TCACGCGGAGCTTACGGTGAACCTTGGCATGGAGACAGTGCATATGCCGAGCTGCC 291
QY 135 ValValAlaLeuTrpValAlaGlnGlyArgGlyTrpHisAlaAsnAsnThrCysAla 154
DB 292 GCACTGATGATGTGGGTGACGAGAAAGTGTGACTACTTCCAAATATGACACATCGCT 351
QY 155 AlAglYArgGlnGlySerGlyThrTrpGlnValAlaTrpArgAsnThrAlaGlyValGly 174
DB 352 GCTGGAACAAACATGCTGATATCTCAAGTGTGTGGAATAAATCAGAGTGGTGA 411
QY 175 CYSAlaAlaIAlaSerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTrpAsnProHis 194
DB 412 TGTGCCAAAGTCAAGTGTGACAAATGGTGAACCTTGTGACTTGCACATATTCATCT 471
QY 195 GlyAsnValGlnGlyInsProtyr 203
DB 472 GGTATGTAGGGAGCTAGCCCTTAC 498

```

## RESULT 7

```

US-09-840-479-6
; Sequence 6, Application US/09840479
; Patent No. US20010025380A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Family of Maize PR-1 Genes and Promoters
; FILE REFERENCE: 5718-32, 035718/175219
; CURRENT APPLICATION NUMBER: US/09/840,479
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/257,583
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 866
; TYPE: DNA

```

; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24

FILE REFERENCE: 1380.001051  
CURRENT APPLICATION NUMBER: US/09/887,576  
CURRENT FILING DATE: 2001-06-25

;  
; APPLICANT: Budworth, P.  
;  
; APPLICANT: Brown, D.  
;  
; APPLICANT: Chang, H.  
;  
; APPLICANT: Zhu, T.  
;  
; APPLICANT: Han, B.  
;

```

: Sequence 85, Application US/
: Patent No. US20020152497A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Rafalski, Antoni
:
: APPLICANT: Miao, Guo-Hua
:

```



```

APPLICANT: Falco, Saverio Carl
APPLICANT: Sakai, Hajime
APPLICANT: Famodu, Omolayo O.
APPLICANT: Odell, Joan T.
APPLICANT: Meyers, Blake
APPLICANT: Thorpe, Catherine
APPLICANT: Weng, Zude
TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
FILE REFERENCE: B1357 US NA
CURRENT FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 09/566,394
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 60/133038
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: 60/133042
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: 60/133427
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133437
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133428
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133438
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133436
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/137667
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: Microsoft Office 97
SEQ ID NO 85
LENGTH: 714
TYPE: DNA
ORGANISM: Oryza sativa
US-10-078-929-85

Alignment Scores:
Pred. No.: 2,92e-17 Length: 714
Score: 287.00 Matches: 70
Percent Similarity: 52.57% Conservative: 22
Best Local Similarity: 40.00% Mismatches: 71
Query Match: 26,45% Indels: 12
Gaps: 7

US-09-832-320-2 (1-203) x US-10-078-929-85 (1-714)
QY 33 AlaProAlaProThrHisGlyAlaArgValLeuMetProGlyGlyAlaGlyAlaValThr 52
Db 22 GCACCTTCGCAAGGTCAAGCCGCGCGCGCTC-----GCCGTGGCATCTCG 69
QY 53 LysAlaInGlnInGlyThrGlySerGlySerAsnAlaThrAlaArgGluTyrLeuAla 72
Db 70 CTGGCCATGGCGGCCACACCCACCTCGGCGCGCAAGACCGCGCATAGCTAAC 129
QY 73 ProHisAnGlnAlaArgAlaValGlyValAlaProLeuArgTyrPasnAlaGlyLeu 92
Db 130 CTGCACAAACAGCGCGCGCGCGCGCGCTCGCGCGAGCTGGAGACCCCAAGCTC 189
QY 93 AlaSerAlaAlaGlyThrValAlaGlnInGlnArgTyrGlnGlyCysAlaPheAla 112
Db 190 GCCAGCTTCGCGCAGACGCTACGCGCC-----AAGCGCGCGCGACTGCGCGCTCAG 243
QY 113 AspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTyrPalaSerTyr--ArgAla--- 130
Db 244 CACTCCGGC---GGCCGTACAGCGCGCAACATCTTGGGGCTCGCGCGCGCGCGCTGG 300
QY 131 ArgProAlaValValAlaLeuTyrValAlaGlnGlyArgTyrTyrThrHisAlaAsn 150
Db 301 AGCGCGCGCAGCGCGGTGGGTGGTGGTGGCGGAGAAAGAACTACCACTAGCAAC 360
QY 151 AsnThrCysAlaAlaGlyArgGlnCysGlyTyrThrGlnValValTyrPArgAsnThr 170
```

```

Db 361 AACACGTGCACCCCGGCAAGGTGTGGCCACACCAACCCAGGTGTGGCCCAAGTGC 420
QY 171 AlaGluValGlyCysAlaGlnAlaSerCysAlaThr-----GlyAlaThrLeuThrLeu 188
Db 421 GTGGCATGTGGGTGGCGCGCGCGCTGTGGCGCGCGGAGAACCGCGCGCTGCATCACC 477
QY 189 CysLeuTyrAsnProHisGlnValGlnGlyInSerProTyr 203
Db 478 TGCAACTACGACCCCGCGGCAACTTCAACGCGAGCGCCGCTTC 522

RESULT 12
US-10-078-929-99
Sequence 99, Application US/10078929
Patent No. US20020152497A1
GENERAL INFORMATION:
APPLICANT: Rafalski, Antoni
APPLICANT: Miao, Guo-Hua
APPLICANT: Falco, Saverio Carl
APPLICANT: Sakai, Hajime
APPLICANT: Famodu, Omolayo O.
APPLICANT: Odell, Joan T.
APPLICANT: Meyers, Blake
APPLICANT: Thorpe, Catherine
APPLICANT: Weng, Zude
TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
FILE REFERENCE: B1357 US NA
CURRENT FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 09/566,394
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 60/133038
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: 60/133042
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: 60/133427
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133437
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133428
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133438
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/133436
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/137667
PRIOR FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: Microsoft Office 97
SEQ ID NO 99
LENGTH: 604
TYPE: DNA
ORGANISM: Triticum aestivum
US-10-078-929-99

Alignment Scores:
Pred. No.: 8,95e-17 Length: 604
Score: 280.50 Matches: 70
Percent Similarity: 48.98% Conservative: 26
Best Local Similarity: 35.71% Mismatches: 61
Query Match: 25,85% Indels: 39
Gaps: 8

US-09-832-320-2 (1-203) x US-10-078-929-99 (1-604)
QY 15 AlaProMetaLThrHisLacCysLeuLeuAlaThrLeuAlaLeuAlaPro 34
Db 12 TCGCGCAAGGTACAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 59
QY 35 AlaProThrHisGlyAlaArgValLeuMetProGlyGlyAlaGlyAlaValThrLysAla 54
Db 60 -----GTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 74
```

```

: LENGTH: 783
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (760)
: US-10-078-929-81

Alignment Scores:
Pred. No.: 3,49e-15 Length: 783
Score: 264.00 Matches: 69
Percent Similarity: 51.41% Conservative: 22
Best Local Similarity: 38.98% Mismatches: 76
Query Match: 24.33% Indels: 10
DB: 12 Gaps: 7

US-09-832-320-2 (1-203) x US-10-078-929-81 (1-783)
OY 32 AAlAAlProAlProThnHnISGLYAlAAlArgValLeuMetProGlyValAGLYAlAVal 51
Db 30 GCGTTCCGAGACCTACTAGTCGTGTACCGCGCTA---GTCGCCCTGCTGCGCCATG 86
OY 52 ThLYsAlAGlnGlnGlnGlyLYThLYGlySerLYSerAlaThrAlAspLIuTYrLeu 71
Db 87 GCGGCGCGATGATGATGCGCGCGCGCGCTCTGCGCGAGAAC---ACGCGCGAGACCTTCGTG 143
OY 72 AAlProHnISAsnGlnAlAArgAlAAlAValGlyLYValAlAProLeuHrGTPrpAsnAGly 91
Db 144 MATCTGCACACACCGCGCGCGCGCGCGGACGCGTGGCGCCGCTGGCGCTGGAGACCCAGG 203
OY 92 LeuAlASerAlaAlAAGLYThrValAlAGlnGlnAArgArgGlnGlnGlyLYsAlAphe 111
Db 204 GTGCGCAGGTAGCGCGAGAGACTACCGCGCG---AAGCGCGCGCGGAGTACGCGGCTG 257
OY 112 AAlAspValGlyAlASerProTYrGlyAlAsnGlnGlyThrAlAserTYr---ArgAla 136
Db 258 GTGCACTCGCGC---GCGCGCTTCGCGCAGACGATCTTCGCGGCTCGCGCGCGCGGCGG 314
OY 131 ---ArgProAlAGLYAlAValAlAAlALeuTYrValAlAGLYAlAArgTYrTYrThHISAla 144
Db 315 TGGAGCGCGCGACCGCGCTGCGGTGCGAGCGAGAGAACATTCACCTGAGC 374
OY 150 AsnAsnThrCysAlAAlAGLYArgGlnCysGlyThrTYrThrGlnAlValAlThrPrArgsn 165
Db 375 AGCAACACCTCGGACCGCGCGAGGTGTGCGGCACATACAGAGAGGTGTGTGGCGAGG 434
OY 170 --ThrAlAGLYAlGlyCysAlAGlnAlASerCysAlAthrngLY---AlAthrLeuHrL 186
Db 435 TGTCCACCGCGACATCGGCTGCGCGCGCTGTCTGCGCGCAACCGCGCGCTTCATCG 494
OY 188 eucYSLeuTYrAsnProHnISGLYAsnValAGlnGlyGlnSerProTYr 203
Db 495 TCTGCACTACGACCGCGCGCGCGCGACATGCACAGCGCGACCGCGCTTC 541

RESULT 14
US-09-840-479-14
: Sequence 14, Application US/09840479
: Patent No. US20010025380A1
: GENERAL INFORMATION:
: APPLICANT: Creane, Virginia
: TITLE OF INVENTION: Family of Maize PR-1 Genes And Promoters
: FILE REFERENCE: 5718-32, 035718/175219
: CURRENT FILING DATE: 2001-04-23
: PRIOR APPLICATION NUMBER: US/09/840,479
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 14
: LENGTH: 806
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (42)..(586)
; NAME/KEY: unsure
; LOCATION: (775)
; OTHER INFORMATION: The nucleotide at this position may be a or g or t
; OTHER INFORMATION: or c.
US-09-840-479-14
```

## Alignment Scores:

Pred. No.:	3,61e-15	Length:	806
Score:	264.00	Matches:	69
Percent Similarity:	51.41%	Conservative:	22
Best Local Similarity:	38.98%	Mismatches:	76
Query Match:	24.33%	Indels:	10
DB:	10	Gaps:	7

US-09-832-320-2 (1-203) x US-09-840-479-14 (1-806)

```
OY 32 AlaAlaProAlaProThiHisGlyAlaArgValLeuMetProGlyGlyAlaGlyAlaVal 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45 GCGTTTCCGAAGCCTACTAGTGTCTAGCCGCGCTA--GCTGCCCTGCTGCGCCATG 101
OY 52 ThrValAlaGlnGlnGlyThrGlySerGlySerAsnAlaThrAlaAspGluTyrLeu 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 GCGGCGCGGATGATGCGCCGACCGCTCGGCGGCAAC--ACGCGCGACGACTTCG 158
OY 72 AlaProHisAsnGlnAlaArgAlaValGlyValAlaProLeuArgTyrAsnAlaGly 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 AMCTGACACACCGCGCGCGCGCGGACGCGGCGCGCGCGCGCGCGCGCGCGCGCG 218
OY 92 LeuAlaSerAlaAlaAlaGlyThrValAlaGlnGlnArgArgGlnGlyGlyCysAlaPhe 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 GTGGCCAGGTACGACGAGGACTACCGCGGCG--AAAGCGCGCGCGGACTCGCGGCTG 272
OY 112 AlaAspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTyrAlaSerTyr--ArgAla 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 GTGACACTCGGGC--GCGCGCTGCGCGAGACATCTTGCGGCTCGCGCGCGCGCGG 329
OY 131 ---ArgProAlaGluValValAlaAlaLeuTyrValAlaGlnGlyArgTyrTyrHisAla 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 TGGAGCCCCCGACGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 389
OY 150 AsnAsnThrCysAlaAlaGlyArgGlnCysGlyTyrTyrThrGlnValValTyrParAsn 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 AGCAACACCTGCGACCCCGCGCAAGGTGTGCGCGCACACACGACGAGGTGTGCGCGCAG 449
OY 170 --ThrAlaGluValGlyCysAlaGlnAlaSerCysAlaThrGly--AlaThrLeuThrL 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 TGTCCACCCCGCATCGGCTCGCGCGCGCGCGCGCTGTCTGCGCGCGACACCGCGGCTTCATCG 509
OY 188 eucysleuTyrAsnProHisGlyAsnValGlnGlyGlnSerProTyr 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 TCTGCACCTACGACCCCGCGCAACGTCACACGACGCGACGCGCGCGCTTC 556

RESULT 15
US-10-078-929-95
; Sequence 95, Application US/10078929
; Patent No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Meng, Jude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; FILE REFERENCE: BB157 US NA
; CURRENT APPLICATION NUMBER: US/10/078, 929
; CURRENT FILING DATE: 2002-02-19
```

```
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Glycine max
US-10-078-929-95
```

Alignment Scores:	1.9e-14	Length:	701
Pred. No.:	255.00	Matches:	53
Score:	53.62%	Conservative:	21
Percent Similarity:	38.41%	Mismatches:	60
Best Local Similarity:	23.50%	Indels:	4
Query Match:	12	Gaps:	4

US-09-832-320-2 (1-203) x US-10-078-929-95 (1-701)

```
OY 66 ThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaValGlyValAlaPro 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 ACCCTCAAGACTCTTCTGATGTGCGACATCGCTGCGAGGTGTGTGTGCTTCA 174
OY 86 LeuArgTyrAsnAlaGlyLeuAlaSerAlaAlaGlyThrValAlaGlnGlnArg 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 CTCTCATGGACACACACCTT--CAAGCTACGCTCAAGGTATGCTCAATGACGAATC 231
OY 106 GlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTyr 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 CCG--GACTGCACCTC--GAACACTCCATGAGACCTTCGCGAGATCTCGCTGAA 285
OY 126 AlaSerTyrArgAlaArgProAlaGluValAlaLeuTyrValAlaGlnGlyArgTyr 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 GGTACGCGGCAATGATGAGGTTCGATGCTGCAAAATTGGCTCACTGAGAACGCTTAC 345
OY 146 TyrThrHisAlaAsnThrCysAlaAlaGlyArgGlnCysGlyTyrTyrThrGlnVal 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 TATGACCACTACTCCACAGCTGTGTGCCATGAT--GAGTGCCTTGCAATTATCACTAGATT 402
OY 166 ValTyrArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThr 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 GTGTGGCGTATCTGTTCACTTGGGTGTGCTAGACCAAGGTATGACCAATGCTGGGTG 462
OY 186 LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlnGlnSerProTyr 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 463 TTGTGTAATTGCACTATTCCTCCACAGCAACATGTGAAGGGACGACCTTAT 516
```

Search completed: February 22, 2003, 04:08:20  
Job time : 114 secs

•  
•  
•  
•



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 22, 2003, 02:00:38 ; Search time 228 Seconds  
(without alignments)  
2005.069 Million cell updates/sec

Title: US-09-832-320-2

Perfect score: 1085

Sequence: 1 MAHSRSHHLLLPAPMATA.....ATLTLCLVPHNGVQGSPPY 203

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-Q=/cgn2.1/USPTO.spool/US09832320/runat\_20022003\_092726.14782/app\_query\_fasta\_1.391  
-DB=N.GeneSeq.101002 -OEW=fastp -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0  
-LOOEXT=0 -UNIT=bits -STRAP=1 -END=1 -MATRIX=blonsum62 -TRANS=human0.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -MORFEXT HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09832320.cgn2.1.391 &runat\_20022003\_092726.14782 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-MARK\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.GeneSeq.101002:\*

- 1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1085	100.0	612	24	ABA96418
2	1085	100.0	898	24	ABA96417
3	348.5	32.1	679	21	AA049084
4	347.5	32.0	683	21	AA037078
5	333	30.7	749	21	AA021207
6	331	30.5	731	16	AA099804
7	322	29.7	846	24	ABL59008
8	304	28.0	679	24	ABL59001
9	304	28.0	771	11	AA006181
10	304	28.0	771	20	AA062806
11	297.5	27.4	779	21	AA047173
12	297.5	27.3	696	11	AA006182
13	296	27.3	696	11	AA006182
14	296	27.3	696	20	AA028807
15	296	27.3	696	20	AA028807
16	295	27.2	721	10	AA091021
17	292.5	27.0	866	20	AA021191
18	288	26.5	664	24	ABL59002
19	288	26.5	860	16	AA099807
20	288	26.5	860	20	AA062839
21	288	26.5	860	20	AA062839
22	287	26.5	1363	10	AA091026
23	286.5	26.4	447	21	AA040410
24	286	26.4	780	21	AA045155
25	286	26.4	809	11	AA006180
26	286	26.4	809	20	AA062805
27	286	26.4	809	20	AA072997
28	286	26.4	2038	10	AA090367
29	286	26.4	2038	20	AA062887
30	286	26.4	2038	20	AA072989
31	285	26.3	2038	11	AA003663
32	282	25.9	772	10	AA091019
33	281.5	25.9	480	13	AA031653
34	277.5	25.6	480	13	AA031652
35	264	24.3	806	20	AA021194
36	260	24.0	2090	10	AA090368
37	260	24.0	2256	20	AA062853
38	260	24.0	2256	20	AA072990
39	259.5	23.9	483	13	AA031654
40	257	23.7	493	21	AA046773
41	256.5	23.6	480	13	AA031656
42	248	22.9	1260	11	AA004693
43	246	22.7	561	21	AA098346
44	244.5	22.5	453	13	AA031655
45	232.5	21.4	940	21	AA052190

## ALIGNMENTS

## RESULT 1

ABA96418 standard; CDNA: 612 BP.

ABA96418:

02-APR-2002 (first entry)

Maize PRI-C10 coding sequence SEQ ID NO 3.

Maize: pathogen-related; PRI-C10; plant; transgenic; gene; ss.

zma mays.

Location/Qualifiers

Key 1.612 /tag= a  
CDS /product= "PRI-C10"

XX US2001049834-A1.  
 XX 06-DEC-2001.  
 XX 10-APR-2001; 2001US-0832320.  
 XX 10-APR-2000; 2000US-195801P.  
 XX (CRAN/) CRANE E H.  
 XX (CRAN/) CRANE V C.  
 XX Crane EH, Crane VC;  
 XX WPI; 2002-121407/16.  
 XX P-PSDB; AAM48742.  
 XX New nucleic acid encoding a pathogen-related protein isolated from  
 PT maize and designated PRI-C10, useful for transforming plants for  
 PT enhanced disease resistance -  
 XX Claim 1; Page 32; 34pp; English.  
 XX The invention relates to an isolated nucleic acid encoding a  
 CC pathogen-related protein PRI-C10. The nucleic acid is used to transform  
 CC plants for enhanced disease resistance.  
 XX Sequence 612 BP; 80 A; 231 C; 239 G; 62 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 3,41e-69 Length: 612  
 Score: 1085.00 Matches: 203  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-832-320-2 (1-203) x ABA96418 (1-612)

OY 1 MetAlaHisSerArgSerHisHisHisLeuLeuLeuProAlaProMetAlaThrAla 20  
 DB 1 ATGGGGCATCTGGCGCAGCACCACCTCTCTGCTCCCGCCCATGGCCAGCGCG 60  
 OY 21 CysLeuLeuLeuAlaThrLeuLeuAlaLeuCysAlaAlaProAlaProThrHisGlyAla 40  
 DB 61 TGTCTGCTCTCTGCGCACCCTCTGCGCTGCGCGCGCGCGCGCGCAGCCAGCGCGCG 120  
 OY 41 ArgValLeuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGly1yThrGly 60  
 DB 121 CGCGTCTCATGCGCGGGCGCGCGCGGTGACCAAGCGAGCGGTGCGCACCGCG 180  
 OY 61 SerGlySerAspAlaThrAlaAspGlyThrLeuAlaProHisAspGlnAlaThrAlaAla 80  
 DB 181 AGCGGACACACGCGAGCGCGAGTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 OY 81 ValGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal 100  
 DB 241 GGGGCGTGGCG 300  
 OY 101 AlaGlnGlnArgArgGlnGlnGlyCysAlaAlaAspAlaGlyAlaSerProTrpGly 120  
 DB 301 GCGGAGCGGCG 360  
 OY 121 AlaAspGlnGlyTrpAlaSerTrpArgAlaArgProAlaGlyAlaValAlaLeuTrpVal 140  
 DB 361 GCGAACCAAGGGGTGGGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
 OY 141 AlaGlnGlyArgGlyTrpThrHisAlaAspThrCysAlaAlaGlyArgGlnGly 160  
 DB 421 GCGGAGCGGCGCTACTACACCGACCAACACAGCGCGCGCGCGCGCGCGCGCGCG 480  
 OY 161 ThrTrpThrGlnValValThrParGAsnThrAlaGlyValGlyCysAlaGlnAlaSerCys 180

DB 481 ACGTACACGACAGTGTGTGGCGCAACACCGCGAGGTGCGCGCAGCGCAGTGC 540  
 OY 181 AlaThrGlyAlaThrLeuThrLeuCysLeuTrpAsnProHisGlyAsnValGlnGlyAla 200  
 DB 541 GCGAGGCGCGCAGCTCAGCTCTGCTGTACACCGCGCGCGCGCGCGCGCGCGCGCG 600  
 OY 201 SerProTrp 203  
 DB 601 AGCCCTTAC 609

RESULT 2  
 ABA96417  
 ID ABA96417 standard; cDNA; 898 BP.  
 AC ABA96417;  
 XX 02-APR-2002 (first entry)  
 DE Maize PRI-C10 encoding cDNA SEQ ID NO 1.  
 XX Maize; pathogen-related; PRI-C10; plant; transgenic; gene; ss.  
 XX Zea mays.  
 XX Key Location/Qualifiers  
 FT 63..674  
 FT /\*\*tag= a  
 FT /product= "PRI-C10"

US2001049834-A1.  
 XX 06-DEC-2001.  
 XX 10-APR-2001; 2001US-0832320.  
 XX 10-APR-2000; 2000US-195801P.  
 XX (CRAN/) CRANE E H.  
 XX (CRAN/) CRANE V C.  
 XX Crane EH, Crane VC;  
 XX WPI; 2002-121407/16.  
 XX P-PSDB; AAM48742.  
 XX New nucleic acid encoding a pathogen-related protein isolated from  
 PT maize and designated PRI-C10, useful for transforming plants for  
 PT enhanced disease resistance -  
 XX Claim 1; Page 30-31; 34pp; English.  
 XX The invention relates to an isolated nucleic acid encoding a  
 CC pathogen-related protein PRI-C10. The nucleic acid is used to transform  
 CC plants for enhanced disease resistance.  
 XX Sequence 898 BP; 161 A; 295 C; 309 G; 133 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 5.13e-69 Length: 898  
 Score: 1085.00 Matches: 203  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-832-320-2 (1-203) x ABA96417 (1-898)

OY 1 MetAlaHisSerArgSerHisHisHisLeuLeuLeuProAlaProMetAlaThrAla 20  
 DB 63 ATGGGGCATCTGGCGCAGCACCACCTCTCTGCTCCCGCCCATGGCCAGCGCG 122  
 OY 21 CysLeuLeuLeuAlaThrLeuLeuAlaLeuCysAlaAlaProAlaProThrHisGlyAla 40

D	123	TGCTTCTCTCTGGCCACCCTCTCTGGCGCTGTGGCGCGCGCGCGCCAGCCACCACGGCGCG	182
Q	41	ArgValLeuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyThrGly	60
D	183	CGCGTCTTCATATCCGGCGCGCGCGCGCGCGCGGTACCACCAAGCGCCAGGTGGCACCGCGC	242
Q	61	SerGlySerAspAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaAla	80
D	243	AGCGGACGACAGCCACCGCGCGGAGCACTACTCTGGCCGCCACACACCGCGCGCGCGCGC	302
Q	81	ValGlyValAlaProLeuAspArgTyrPasnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal	100
D	303	GTGGGCGGTGGCCCCCTCGGTGGTGAACGCGGGCGCTGGCTTCGCGCGCGCGCGGGACGGTG	362
Q	101	AlaGlnGlnArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly	120
D	363	GCGCACACACGGCGCGCGCGCGGGGTGGCGGTGGCGGACGTGGGGGCCACGCCCTACGGC	422
Q	121	AlaAsnGlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaLeuThrVal	140
D	423	GCCAAACACAGGGGTGGGCGACGTACCGCGCGCGCGCGCGGACGTGGTGGCGCTGGGTG	482
Q	141	AlaGlnGlyArgTyrTyrThrHisAlaAsnAspThrCysAlaAlaGlyArgGlyCysGly	160
D	483	GCGGAGGGCGGTACTATACACCCAGCCACACACAGTGGCGCGCGGGCGCGGCACTGGCGC	542
Q	161	ThrTyrThrGlnValAlaTyrArgAspThrAlaGluValGlyCysAlaGlnAlaSerCys	180
D	543	ACGTACACACAGGTGGTGGCGCACACCGCGGAGGTGGGTGGCGCGCGCGACGTGC	602
Q	181	AlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGln	200
D	603	GCCACGGGCGCGCACGCTGCACCTGTGGCTGTACACCCCGACGGCACAGGTGACAGGCGCAG	662
Q	201	SerProTyr 203	
D	663	AGCCCTTAC 671	
RESULT 3			
AA	49084	ID	AA
AC	49084	standard; DNA; 679 BP.	
AC	49084;		
AC	18-OCT-2000	(first entry)	
XX	18-OCT-2000		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 59865.		
XX	Arabidopsis thaliana		
XX	Hybridisation assay; genetic mapping; gene expression control;		
XX	Protein identification; signal transduction pathway;		
XX	metabolic pathway; promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
XX	03-MAR-1999; 99US-0123180.		
XX	09-MAR-1999; 99US-0123548.		
XX	23-MAR-1999; 99US-0125788.		
XX	25-MAR-1999; 99US-0126264.		
XX	29-MAR-1999; 99US-0126785.		
XX	01-APR-1999; 99US-0127462.		
XX	06-APR-1999; 99US-0128234.		
XX	08-APR-1999; 99US-0128714.		
XX	16-APR-1999; 99US-0129845.		
XX	19-APR-1999; 99US-0130077.		
XX	21-APR-1999; 99US-0130449.		
XX	23-APR-1999; 99US-0130510.		

PR	23-APR-1999	9905-0130891
PR	28-APR-1999	9905-0131449
PR	30-APR-1999	9905-0132048
PR	04-MAY-1999	9905-0132407
PR	04-MAY-1999	9905-0132484
PR	05-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132486
PR	07-MAY-1999	9905-0132487
PR	11-MAY-1999	9905-0132456
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134421
PR	14-MAY-1999	9905-0134721
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134941
PR	20-MAY-1999	9905-0135124
PR	24-MAY-1999	9905-0135353
PR	25-MAY-1999	9905-0135629
PR	27-MAY-1999	9905-0136021
PR	28-MAY-1999	9905-0136592
PR	01-JUN-1999	9905-0136782
PR	03-JUN-1999	9905-0137222
PR	04-JUN-1999	9905-0137528
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	14-JUN-1999	9905-0138819
PR	14-JUN-1999	9905-0139111
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139617
PR	22-JUN-1999	9905-0139617
PR	22-JUN-1999	9905-0139639
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140655
PR	26-JUN-1999	9905-0140823
PR	26-JUN-1999	9905-0140823
PR	26-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141284
PR	01-JUL-1999	9905-0141542
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	12-JUL-1999	9905-0142877
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143524
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144632

PR	0-JUL-1999;	9905-0144884;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145951;
PR	28-JUL-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	05-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-01493175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-01541779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155487;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159294;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159332;
PR	14-OCT-1999;	9905-0159333;
PR	14-OCT-1999;	9905-0159337;

PR	18-OCT-1999	9905-01596384
PR	18-OCT-1999	9905-01595884
PR	21-OCT-1999	9905-01607471
PR	21-OCT-1999	9905-01607641
PR	21-OCT-1999	9905-01607658
PR	21-OCT-1999	9905-01607768
PR	21-OCT-1999	9905-01607704
PR	21-OCT-1999	9905-01608174
PR	21-OCT-1999	9905-01608154
PR	21-OCT-1999	9905-01608155
PR	22-OCT-1999	9905-01609880
PR	22-OCT-1999	9905-01609881
PR	22-OCT-1999	9905-01609889
PR	25-OCT-1999	9905-01614004
PR	25-OCT-1999	9905-01614005
PR	25-OCT-1999	9905-01614006
PR	25-OCT-1999	9905-01614007
PR	26-OCT-1999	9905-01613650
PR	26-OCT-1999	9905-01613659
PR	26-OCT-1999	9905-01613660
PR	28-OCT-1999	9905-01619321
PR	28-OCT-1999	9905-01619320
PR	28-OCT-1999	9905-01619322
PR	28-OCT-1999	9905-01619393
PR	29-OCT-1999	9905-01621442

**Alignment Scores:**

pred. NO.:	1,42e-16	length:	679
Score:	348.50	Matches:	69
Percent Similarity:	60.14%	Conservative:	17
Best Local Similarity:	48.25%	Mismatches:	54
Query Match:	32.12%	Indels:	3
DB:	21	Gaps:	2

US-09-832-320-2 (1-203) x AAC49084 (1-679)

Oy	64	AsnAlaThrAlaAarGluTyrLeuAlaIleProHisAsnGlnAlaIrrAlaValGlyVal	83
Db	201	TCTGGAGCAGCTAAAGCTTTCCACCGATGCCACAAAGGCCAGGCATGGTGTTGT	260
Oy	84	AlaProlaArgTyrPasnAlaGlyLeuAlaSerAlaAlaIaGlyThrValAlaGlnGln	103
Db	261	CCACCACACTAGTTTGGAGCCAGCAACGTTGGAAAGCTGTCGGAAGTCGGTTGGCTGTTACAG	320
Oy	104	ArgArGargInGlyGlyCysAlaIlePheAlaAspValGlyAlaSerProTyrGlyAlaAsnGln	123
Db	321	AGGAACCAAAAAGAAGTGTGAGTGGCGCGAGTCTAAACCCCTGGAAAAATACGGCGCAACCG	380
Oy	124	GlyTrrAlaSer-----TyrArgAlaIrrProAlaGluValAlaIalaLeuTrpAlaLa	141
Db	381	CTTTGGGGCTAAAGCGCTTAGTACCCGTGACACCGCTCTGCTGGTGGAGACTGGGTAAG	440
Oy	142	GluGlyArgTyrTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThr	161
Db	441	GAGAAACCTTTCTACAATTATTAAGTCACACACAGTGTGCTGGCAACCCACAGCTGGGGGTT	500
Oy	162	TyrThrGlnValValTrrPrArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAla	181
Db	501	TATAAACAAAGTGTGTGGAGAAACCTTAAAGCGCTGGGTGTGCTCAAGCACAGCTGTACG	560
Oy	182	ThrGlyAlaThr---LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGln	200
Db	561	AAGAGTCAACGCGTGTGACCATTTGTTTTCACAATCTCTCGTGAATAGTAAATGGCCAA	620
Oy	201	SerProTyr	203
Db	621	AAGCTTAC	629
<hr/>			
RESULT 4			
AAC37078			
ID	AAC37078 standard; DNA; 683 bp.		
XX			
AC	AAC37078;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 16099.		



```
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0148722.
PR 20-AUG-1999; 99US-0148723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
```

```

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

## Alignment Scores:

```

Pred. No.: 1.69e-16 Length: 683
Score: 347.50 Matches: 68
Percent Similarity: 60.14% Conservative: 18
Best Local Similarity: 47.55% Mismatches: 54
Query Match: 32.03% Indels: 3
DB: 21 Gaps: 2

```

US-09-832-320-2 (1-203) x AAC37078 (1-683)

```

OY 64 AsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaValGlyAla 83
DB 201 TGTGAGAGCAAGCTTACCGATGCGCAGACGACGACGACGACGACGACGACGACG 260
OY 84 AlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAlaGln 103
DB 261 CCACCACTAGTTGGAGCCAGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 320
OY 104 ArgArgGlnGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaGln 123
DB 321 AGGAACCAAAAGAGTGAAGTGCAGAGTCTAAACCCCTGGAAGAAAGAGCGGAGACG 380
OY 124 GlyTPAlaSer-----TyrArgAlaArgProAlaGluValAlaLeuTyrValAla 141
DB 381 CTTTGGCTAAGGGCTTAGTAGCGGTGACACCGCTCTCTCTGTCGTGAGACTTGGTGAAG 440

```

```

OY 142 GluGlyArgTyrTyrThrHisAlaAsnThrCysAlaAlaGlyArgGlnCysGlyThr 161
DB 441 GAGAAACCTTCTCAATATATAGTCAGACGCTGCTGCGACACGACGACGACGACGCTT 500
OY 162 TrrThrGlnValAlaTrrArgAsnThrAlaGlnValGlyCysAlaGlnAlaSerCysAla 181
DB 501 TATTAACAGTCTGTGAGAACTCTAAAGACCTGCGGTGCTCCAGCCAGCTTCG 560
OY 182 ThrGlyAlaThr---LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGln 200
DB 561 AAAGAGTCAGCGTGTGACCATTTGTTTACAACTCTCTGGAATATATATATGCGCAA 620
OY 201 SerProTyr 203
DB 621 AAGCCTTAC 629

RESULT 5
AAZ21207 standard; DNA; 749 BP.
ID AAZ21207
XX
AC AAZ21207;
XX
DE 22-NOV-1999 (first entry)
XX
XX Zea mays pathogenesis-related class I PR-1#83 gene.
XX
XX Zea mays; maize; pathogenesis-related class I; PR-1; promoter;
XX regulation; expression; disease resistance; genetic manipulation;
XX tobacco mosaic virus; cucumber mosaic virus; ringspot virus;
XX necrosis virus; maize dwarf virus; viroid; bacterial; insect;
XX nematode; fungal; ss.
XX
OS Zea mays.
XX
FH Key 92..583 Location/Qualifiers
FT CDS /*tag= a
FT /*product= "PR-1#83"
FT /*note= "pathogenesis-related class I protein"
XX
XX WO9943819-A1.
XX
PD 02-SEP-1999.
XX
XX 11-FEB-1999; 99WO-US03011.
XX
XX 26-FEB-1998; 98US-0076100.
XX 27-MAR-1998; 98US-0079648.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Crane VC;
XX
XX WPI: 1999-527621/44.
XX P-PSDB: AAZ29944.
XX
XX New promoter sequences from pathogenesis-related genes of maize
XX Example 3; Page 72-73; 86pp; English.
XX
XX AAZ21186 to AAZ21190 represents the nucleotide sequences for promoters
XX isolated from a family of maize (Zea mays) genes encoding pathogenesis
XX related (PR-1) proteins. The promoters are useful for expressing
XX heterologous genes (including genes for disease resistance) in plants,
XX especially dicots, or monocots i.e. maize. The promoters are useful for
XX the genetic manipulation of plants to exhibit specific phenotypes,
XX particularly enhanced resistance to pathogen-caused disease. Pathogens
XX include viruses such as tobacco or cucumber mosaic virus, ringspot
XX virus, necrosis virus, and maize dwarf virus, and viroids, bacteria,
XX insects, nematodes and fungi. The present sequence encodes a maize
XX PR-1 protein given in the present invention.
XX

```

Sequence 749 BP: 160 A; 235 C; 206 G; 148 T; 0 other;

## Alignment Scores:

Pred. No.:	2,02e-15	Length:	749
Score:	333.00	Matches:	81
Percent Similarity:	48.13%	Conservative:	22
Best Local Similarity:	37.85%	Mismatches:	67
Query Match:	30.69%	Indels:	44
DB:	20	Gaps:	8

US-09-832-320-2 (1-203) x AA099804 (1-749)

```

QY 5 ArgSerHisHisLeuLeuLeuProAla-----
DB 26 CGATCAGACATTGACCTTGCCTGCTCATATTAATAGTTCATCAGCAACAA 85
QY 16 ---Prometala-----ThralaCysLeuLeuAlaThrlaLeuAlaLeuCysAla 32
DB 86 CAACAGATGGCAGCAGGCTAGCGCTCCTAGCTGCGCATGAGCAGCATGCTGTG 145
QY 33 AlAProAlaProThrlaHisGlyAlaAlaValLeuMetProGlyAlaGlyAlaValThr 52
DB 146 GCGCCGTGC-----
QY 53 LysAlaGlnGlnGlyThrlGlySerGlySerAlaAlaThrlaAlaAspGlyLeuAla 72
DB 155 -----ACGGCCAGAACTCGCCGAGAAC---TACGTGGAC 167
QY 73 ProHisAsnGlnAlaAlaAlaAlaValGlyAlaProLeuAlaGlyTrpAsnAlaGlyLeu 92
DB 188 CCGCAGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247
QY 93 AlAserAlaAlaAlaGlyThrlValAlaGlnGlnAlaGlnGlnGlnGlnGlnGln 112
DB 248 GCGCGGTAGCGCAGACGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
QY 113 AspValGlyAlaSerProGlyAlaAlaAsnGlnGlyTrpAlaSerTrpAlaGlyAla 130
DB 302 CACTCCGGC---GGCCCTACGCGCGCAGAACTTCTGCGCGCGCGCGCGCGCG 358
QY 131 ArgProAlaGlnValAlaAlaLeuTrpValAlaGlnGlnGlnGlnGlnGlnGln 150
DB 359 TCGGGGTCCGAGCGCGCTGCGGTCTGCGGTCTCGAAGACAGTCTAGCAGCAGAC 418
QY 151 AsnTrpCysAlaAlaGlyAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 170
DB 419 AACAGCTGGCGGAGGCGCGGTGCGCGCAGTACAGCAGGAGTGTGCGCGCGCTCC 478
QY 171 AlAGluValGlyCysAlaGlnAlaSerCysAlaThrlGlyAla---ThrlLeuThrLeuCys 189
DB 479 ACCGGCATGCGCTGTCGCCGCGCTGTCGCGAACAACAGCGCGCGCTTCATCATCTGC 538
QY 190 LeuTrpAsnProHisGlyAlaAsnValGlnGlnGlnGlnGlnGlnGlnGlnGln 203
DB 539 ACCTACAAACCGCGCGCGCAGCTGCTGCGGAGAGCCCTTAC 580

RESULT 6
AA099804 ID AA099804 standard; cDNA; 731 BP.
XX AC AA099804;
XX AC
XX 20-JUN-1996 (first entry)
XX
XX PR-1 like gene PR-1mz.
XX
XX SAR: tobacco; protein-synthesis independent gene; cycloheximide;
XX systemic acquired resistance response; anti-pathogen; plant protection;
XX maize; PR-1; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
FH

```

FT CDS 40..531  
FT /\*tag= a

PN W09519443-A2.

PD 20-JUL-1995.

PF 03-JAN-1995; 95WO-IB00002.

PR 13-JAN-1994; 94US-0181271.

PA (CIBA ) CIBA GEIGY AG.

PI Alexander DC, Ryals JA, Uknes SJ, Ward ER;

DR WPI, 1995-263872/34.

PS P-PSDB; AAR91595.

PT New DNA contg. plant systemic acquired resistance genes - and  
transgenic plants contg. them, impart disease and pest resistance,  
also Arabidopsis gene promoter to control DNA transcription

Claim 21; Page 70; 85pp; English.

This sequence represents the DNA sequence of a maize PR-1 like gene.  
CC PR-1mz. This sequence was isolated by screening a BTH-induced cDNA  
CC library of maize. The library was screened using a probe matching to  
CC the PR-1 barley clone HVPRI1R. This sequence, AA099800-099803 and  
CC AA099805 are all used in recombinant/chimeric DNA molecules of the  
CC invention. These sequences were isolated by differential screening of a  
CC cDNA library, followed by analysis by Northern hybridisation to RNA in  
CC the presence and absence of cycloheximide. The genes are used in the  
CC creation of transgenic plants. All of these sequences confer  
CC anti-pathogenic properties to transgenic plants. Transgenic expression  
CC of 2 or more of the recombinant molecules of the invention that encode  
CC anti-pathogenic proteins provides a synergistic increase in plant  
CC protection, and may also offer protection against a wider range of  
CC pathogens.

Sequence 731 BP: 186 A; 212 C; 199 G; 134 T; 0 other;

## Alignment Scores:

Pred. No.:	2,74e-15	Length:	731
Score:	331.00	Matches:	73
Percent Similarity:	50.27%	Conservative:	21
Best Local Similarity:	39.04%	Mismatches:	32
Query Match:	30.51%	Indels:	6
DB:	16	Gaps:	6

US-09-832-320-2 (1-203) x AA099804 (1-731)

```

QY 20 AlAcysLeuLeuLeuAlaThrlLeuAlaLeuCysAlaAlaProAlaProThrlHisGly 39
DB 55 GCGTCTCTCTTACGCTGCGCATGCGCATCTGCTGCGCGCGCTGC----- 102
QY 40 AlaArgValLeuMetProGlyGlyAlaGlyAlaValThrlValAlaGlnGlnGlyThr 59
DB 102 -----
QY 60 GlySerGlySerAlaAlaThrlAlaAspGlyTrpLeuAlaProHisAsnGlnAlaAlaGly 79
DB 103 ---ACGGCCAGAACTCGCCGAGAAC---TACGTGAGAACCGCAGAACGCGCGCGCC 156
QY 80 AlAValGlyValAlaProLeuAlaGlyTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThr 99
DB 157 GACGTGGGCGTGGCGCGGTCTGCGGACACACACGCTGGCGGTGACGCGCAGACTAC 216
QY 100 ValAlaGlnGlnAlaGlyAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 119
DB 217 GCGGCGCAG-----CGCAGGCGCAGCAGCAAGCTGATCCACTCGCGC---GGGCGCTAC 267
QY 120 GlyAlaAsnGlnGlyTrpAlaSerTrpAlaGlyAla-----ArgProAlaGlnValAlaAla 137

```

```

Db      268 GCGGAAACCTCTTCTGGGGCTCCGCGCGCACTGTGTGGGCTCCGACGCCGTGGG 327
Qy      138 LeuTPVAlAlaIgluGlyArgTyrThrHisAlaAsnPrpCysAlaIglValArg 157
      |||||.....:|||||  |||  |||||.....:|||||  |||||.....:
Db      328 TCCTGGGTGTCTGCAAGACAGTACTACACACACACACACGCGGAGGGGAG 387
Qy      158 GlnCysGlyThrTyrThrGlnValValTPrpArgAsnThrAlaGlnValGlyCysAlaGln 177
      |||||  |||||.....:|||||  |||||.....:  |||||.....:
Db      388 GTGTGCGCGCACTACACGCGAGGTGTGGCGCGCATCCGCCGCTGTGGCCGC 447
Qy      178 AlaSerCysAlaThrGlyAla---ThrLeuThrLeuCysLeuTyrAsnPrpHisGlyAsn 196
      |||  |||  |||||.....:|||||  |||||.....:
Db      448 GTGCTGTGCGACACACACGCGCGCTTTCATCTGCACTGACCTACACCGCGGGGAC 507
Qy      197 ValGlnGlyGlnSerProTyr 203
      |||  |||||.....:|||||  |||||.....:
Db      508 GTGCTGCGCGAGAGCCCTAC 528

RESULT 7
ABL59008
ID      ABL59008 standard; DNA; 846 BP.
XX
AC      ABL59008;
XX
DT      20-AUG-2002 (first entry)
XX
DE      Nucleotide sequence of an antibacterial protein.
XX
KW      Antibacterial protein; microbe resistance; plant; gene; ss.
XX
OS      Elaeis guineensis.
XX
XX
FH      Key      Location/Qualifiers
FT      CDS      35..523
FT      /tag= a
FT      /product= "antibacterial protein"
FT
FT
PA      JP2002095477-A.
XX
XX
PD      02-APR-2002.
XX
PF      20-SEP-2000; 2000JP-0285905.
XX
PR      20-SEP-2000; 2000JP-0285905.
XX
XX
PA      (MITU ) MITSUBISHI CHEM CORP.
XX
PA      (BADA-) BADAN PENGKAJIAN DAN PENKERAPAN TEKNOLOGI.
XX
PA      (PAKR-) PT PAKRIE BROS.
XX
PA      (BIOI-) BIOINDUSTRI KYOKAI SH.
XX
PA      (DOKU-) DOKURITSU GYOSSEI HOJIN SANGYO GIJUTSU SO.
XX
DR      WPI; 2002-439987/47.
XX
DR      P-PSDB; ABB77767.
XX
XX
PT      New protein and its gene, useful for creating plants with high
XX
PT      resistance to pathogenic microbes -
XX
PS      Claim 3; Page 8-9; 13pp; Japanese.
XX
XX
CC      The present sequence encodes a polypeptide which has antibacterial
CC      activity. The antibacterial protein and its polynucleotide can be used
CC      for the creation of a plant with resistance against pathogenic microbes.
XX
SQ      Sequence 846 BP; 218 A; 208 C; 214 G; 206 T; 0 other;

Alignment Scores:
Pred. No.:      1..4e-14      Length:      846
Score:          322.00      Matches:      62
Percent Similarity: 61.76%      Conservative: 22
Best Local Similarity: 45.59%      Mismatches: 48
Query Match:    29.68%      Indels:      4
DB:            24      Gaps:      3

```

```

US-09-832-320-2 (1-203) x ABL59008 (1-846)
Qy      69 GlnTyrLeuAlaIleProHisAsnGlnAlaArgAlaValGlyValAlaProLeuArgTyr 88
      |||||.....:|||||  |||||.....:|||||  |||||.....:
Db      122 GACTTGTGACCGCCCAACACGACCGCCGCGCTGCGCGCGCGCGGTGTGG 181
Qy      89 AsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAlaGlnGlnArgArgGlnGly 108
      |||||.....:|||||  |||||.....:
Db      182 GACAAACACCGTGGGAGCGCTACCGCCAGACCTACGCCAACAG-----CGAATCGGCGAC 235
Qy      109 CysAlaIleAlaIleAlaValGlyAlaSerProTyrGlyAlaAsnGlnGlyThrAlaSerTyr 128
      |||||.....:|||||  |||||.....:
Db      236 TGCAGCTGTGCTACCTCCGGC---GGACCGTACGCTAGAACCTCTTGTGGGATAGGT 292
Qy      129 ArgAla---ArgProAlaGlnValAlaIleThrValAlaGlnGlyArgTyrTyrThr 147
      |||||.....:|||||  |||||.....:
Db      293 AGGAGTACACCGACGCGAGATGCTGTGAACCTGTGGGTGATGAGAACGAGTGTGAC 352
Qy      148 HisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyrThrGlnValValTPrp 167
      |||||.....:|||||  |||||.....:
Db      353 TACACGACCAACACTGTGGCGCGGAGGTGTGGCTACTACACACCGAGTGTGTGG 412
Qy      168 ArgAsnThrAlaGlnValGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThrLeuThr 187
      |||||.....:|||||  |||||.....:
Db      413 CGAGACTCCACCCACATCGGTGGCTCGGGGTGAATGCAACAGTGGCGCATCTTATT 472
Qy      188 LeuCysLeuTyrAsnPrpHisGlyAsnValGlnGlyClnSerProTyr 203
      |||||.....:|||||  |||||.....:
Db      473 ATCTGCACACTACAAACCTCCGCGCAATATCGTGGCGACGCCCATAT 520

RESULT 8
ABL59001
ID      ABL59001 standard; DNA; 679 BP.
XX
AC      ABL59001;
XX
DT      20-AUG-2002 (first entry)
XX
DE      Nucleotide sequence of an antibacterial protein.
XX
KW      Antibacterial protein; antibacterial agent; food; plant; gene; ss.
XX
XX
OS      Wasabia japonica.
XX
XX
FH      Key      Location/Qualifiers
FT      CDS      44..529
FT      /tag= a
FT      /product= "antibacterial protein"
FT
FT
PA      JP2002095475-A.
XX
XX
PD      02-APR-2002.
XX
PF      19-SEP-2000; 2000JP-0284178.
XX
PR      19-SEP-2000; 2000JP-0284178.
XX
XX
PA      (IWAT-) IWATE KEN.
XX
XX
DR      WPI; 2002-439986/47.
XX
DR      P-PSDB; ABB77765.
XX
XX
PT      New antibacterial protein gene of Wasabia japonica -
XX
PS      Claim 3; Page 12-13; 17pp; Japanese.
XX
XX
CC      The present sequence encodes an antibacterial protein of Wasabia
CC      japonica. The protein can be used in an antibacterial agent and a
CC      functional food.
XX
SQ      Sequence 679 BP; 211 A; 142 C; 156 G; 170 T; 0 other;

Alignment Scores:

```

Pred. No.: 2,14e-13 Length: 679  
 Score: 304.00 Matches: 72  
 Percent Similarity: 50.85% Conservative: 20  
 Best Local Similarity: 39.78% Mismatches: 61  
 Query Match: 28.02% Indels: 28  
 Gaps: 7

US-09-832-320-2 (1-203) x ABL59001 (1-679)

QY 23 LeuLeuAlaThrLeuLeuAlaLeuCysAlaAlaProAlaProThrHisGlyAlaAlaGlyVal 42  
 DB 68 CTTCCTAAATCTTGGACCCCTGTAGAGCT-----ATT 103  
 QY 43 LeuMetProGlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyGlyThrGlySerGly 62  
 DB 104 GTCTCTCC-----TCGAGGCTCAAGAC----- 127  
 QY 63 SerAsnAlaThrAlaAspGluThrLeuAlaProHisAsnGlnAlaArgAlaAlaValGly 82  
 DB 128 -----AGCCACAAAGACTATCTAAGGTTCCACACGACGACGCGGTAGGC 178  
 QY 83 ValAlaProLeuAlaArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAlaGln 102  
 DB 179 GTTGGCCCAATGCGAGTGGACGACGAGGTTCCA--GCCCTTCGCTCGGAGCTACGCGAC 235  
 QY 103 GlnAlaArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsn 122  
 DB 236 CAACGAAAGA---GGCAGCTCCAGGCTCATACCTGTGT---GGTCTTACGGCGAGAAC 289  
 QY 123 GlnGlyTrpAlaSerTyrTrpAlaArgProAlaGluValValAlaLeuTrpValAlaGlu 142  
 DB 290 TTGGCCCTGGGGAAGCAGCAGCTTATCTGGCATMACCGCCGAGAACATGTGGGTAAACGAG 349  
 QY 143 GlyAlaArgTyrThrHisAlaAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyr 162  
 DB 350 AAGGCTACACAACTACACCTTCGAACACGTGC-----AATGAGATTGTGTGCTACTAC 403  
 QY 163 ThrGlnValValTrpArgAsnThrAlaGluValGlyCysAlaGlnAlaAlaSerCysAlaThr 182  
 DB 404 ACTGAGTGTGTGGAGAACTCGGTGAGACCTGGATGCTAAAGTAGAGGTAAACAAAT 463  
 QY 183 GlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGlnSerPro 202  
 DB 464 GGTGGAACCATCATCGTTGGCAACTATGATCTCTCTGGCAATTAATGGAACCAAAACCT 523  
 QY 203 Tyr 203  
 DB 524 TAC 526  
 RESULT 9  
 AAQ06181 standard; DNA; 771 BP.  
 AC AAQ06181:  
 CT 31-JAN-1991 (first entry)  
 XX PR-1b cDNA cloned into plasmid pBSPR1-1023.  
 DE Transgenic plants; disease resistance; chimeric DNA;  
 KW plant pathogenesis-related protein; ss.  
 OS synthetic.  
 XX Key Location/Qualifiers  
 FH mat\_peptide 120..533  
 FT /tag a  
 FT /product-PR-1b  
 XX EP392225-A.  
 XX PD 17-OCT-1990.

PF 21-MAR-1990; 90EP-0105336.  
 XX 20-OCT-1989; 89US-0425504.  
 PR 24-MAR-1989; 89US-0329018.  
 PR 20-JUN-1989; 89US-0368672.  
 XX  
 PA (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Rials JA, Alexander DC, Goodman RM, Meins F, Payne GB;  
 PI Stinson JR, Neuhaus J-W, Moyer MB;  
 XX  
 DR WPI; 1990-313983/42.  
 DR P-PSDB; AAR07314.  
 XX  
 PT Disease-resistant transgenic plants - obtd. using encoding an  
 PT Inducible pathogenesis-related protein from infected plants.  
 XX  
 PS Example 15; page 24; 77pp; English.

CC This chimeric DNA sequence comprises the PR-1b plant pathogenesis-  
 CC related protein (PRP)-coding sequence (I), and a promoter sequence  
 CC which enhances transcription of (I). This construct is used to  
 CC produce transgenic plant cells or -tissues with the ability to reg-  
 CC erate into plants which are disease resistant.  
 CC See also AAQ06179-80, AAQ06182-86, AAQ06199-006208 and AAQ06829.  
 XX  
 SQ Sequence 771 BP; 227 A; 148 C; 156 G; 240 T; 0 other:

#### Alignment Scores:

Pred. No.: 2,45e-13 Length: 771  
 Score: 304.00 Matches: 64  
 Percent Similarity: 56.55% Conservative: 18  
 Best Local Similarity: 44.14% Mismatches: 57  
 Query Match: 28.02% Indels: 6  
 Gaps: 3

US-09-832-320-2 (1-203) x AAQ06181 (1-771)

QY 61 SerGlySerAsnAlaThrAlaAspGluThrLeuAlaProHisAsnGlnAlaArgAlaAla 80  
 DB 111 TCTCATGCCCAAACTCATACAAAGACTATTTGGTGGCTTACACAGCTCGTCGAGAT 170  
 QY 81 ValGlyValAlaProLeuAlaArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal 100  
 DB 171 GTAGCGGTGGAACCATTAACCTTGGGACAGCGGGTACAGCTTGCACAAATTATGTT 230  
 QY 101 AlaGlnAlaArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120  
 DB 231 TCTCAATTGGCTCGACAGCTGCAACCTCGTACATTTCTCAT-----GGCCAATCGCG 281  
 QY 121 AlaAsn-----GlnGlyTrpAlaSerTyrArgAlaArgProAlaGluValAlaLeu 138  
 DB 282 GAAMACCTAGCTCAGGGAAGTGGCGATTTTATGAG--GCTGTAGAGCGCTCGAGAT 338  
 QY 139 TrpValAlaGluGlyArgTyrThrHisAlaAsnThrCysAlaAlaAlaGlyArgGln 158  
 DB 339 TGGGTGATGAAAGAACGATCTATGACCTCAAAATCTTGTGCAAGACAGAGGTG 398  
 QY 159 CysGlyThrTyrThrGlnValAlaTrpArgAsnThrAlaGluValGlyCysAlaGlnAla 178  
 DB 399 TGTGACACATATCTCAGGTGGTGGCTTACCTGCTGCTGATGTGCTAGAGTT 458  
 QY 179 SerCysAlaThrGlyValaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
 DB 459 AAGTGCACAAATGAGGATATGTGTCTTTCACACTATGATCTCCAGGTATGTGATA 518  
 QY 199 GlyGlnSerProTyr 203  
 DB 519 GGCCAAAAGCTCATAC 533  
 RESULT 10  
 AAQ62806 standard; cDNA; 771 BP.  
 ID AAQ62806

```

XX AAV62806;
AC
XX 05-MAR-1999 (first entry)
DT
XX Tobacco PR-1b gene clone.
DE
XX Chemically regulatable DNA promoter; expression control; pesticide;
KW herbicide tolerance; pathogenesis related gene; PR gene; ss.
XX Nicotiana acuminata.
XX US5851766-A.
XX 22-DEC-1998.
XX 31-MAY-1995; 95US-0456262.
XX 31-MAY-1995; 95US-0456262.
XX 31-MAY-1995; 95US-0456262.
XX (NOVS ) NOVARTIS FINANCE CORP.
XX Harms C, Ryals JA;
XX WPI: 1999-080396/07.
XX Isolating chemically regulatable DNA sequences in plants - useful
PT for chemically controlling expression in transformed plants
XX
XX Example 44; Column 181-182; 175pp; English.
XX
XX This sequence represents a clone of the tobacco pathogenesis related (PR)
CC gene. This gene can be isolated using the method of the invention.
CC The method is for isolating a chemically regulatable DNA promoter
CC fragment from the 5' flanking region of a chemically regulatable gene in
CC a plant tissue. The method allows isolation of sequences which will be
CC useful for the controlled expression of genes, under the control of a
CC non-coding regulatable sequence. This is useful in plants with a
CC herbicide or pesticide detoxification mechanism under the control of a
CC chemical regulator, the regulator being applied before or with the
CC herbicide or pesticide to give optimal tolerance. The promoter fragment
CC is useful for controlling sequences which encode traits such as
CC height, shape, development, male or female sterility, and the ability
CC of the plant to withstand cold, heat, salt and drought. The chemical
CC induction of the promoter allows the regulation of production of
CC compounds, e.g. flavours, fragrances, pigments, natural sweeteners,
CC industrial feedstocks, antimicrobials and pharmaceuticals, by
CC biosynthesis or metabolite conversion, whose biosynthesis is controlled
CC by endogenous or foreign genes. The method allows control over the time
CC and rate of gene expression either throughout the whole plant, or in
CC localized tissues, to achieve e.g. fungal or insect resistance by for
CC instance dusting the leaves with the chemical regulator. Controlling the
CC developmental processes by the application of a regulating chemical in,
CC e.g. the commercial production of cultivated crops allows processes such
CC as germination, flower formation and fruit ripening to be synchronised at
CC a given time.
XX
SQ Sequence 771 bp; 227 A; 148 C; 156 G; 240 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 2.45e-13 Length: 771
Score: 304.00 Matches: 64
Percent Similarity: 56.55% Conservative: 18
Best Local Similarity: 44.14% Mismatches: 57
Query Match: 28.02% Indels: 6
DB: Gaps: 3

```

US-09-832-320-2 (1-203) x AAV62806 (1-771)

```

OY 61 SerGlySerAsnAlaIhrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaAla 80
||| :::: ::::| ||||| ||||| ||||| |||||
DB 111 TCTCATGCCCAAACTCTCAACAGACTATTGTGATGCCATTAACACAGCTGTGCAGAT 170

```

```

OY 81 ValGlyValAlaProLeuArgTyrPAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 171 GTAGCGGTGAGAACCTTAACTTGGACACAGGGGTGACAGCCTATGCAAAATATTATGTT 230
OY 101 AlaGlnGlnArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGly 120
::: ||| ::::| ||||| ||||| ||||| ||||| ||||| |||||
DB 231 TCTCAATTGGCTGCAGACTCAACCTGATTCATTTTCAT-----GCCCAATACGCG 281
OY 121 AlaAsn-----GlnGlyTyrPalasetrYArgAlaArgProAlaGluValAlaLeu 138
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 282 GAACACCTAGCTCAGGAGAAAGTGGCGATTATGACG---GCTGCTAAGCGCTGACAGT 338
OY 139 TrpValAlaGlnArgTyrTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGln 158
||||| ||| ::::| ||||| ||||| ||||| ||||| ||||| |||||
DB 339 TGGTGCATGAGAAACAGTACTATGACATGACATAAATCTTGTCACAAAGACAGCGTG 398
OY 159 CysGlyThrTyrThrGlnValValTTPArgAsnThrAlaGluValGlyCysAlaGlnAla 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 399 TGTGACACTTACTCAGCGGTGTTGGCGTAACTCCGTTCCGTTGATGATGCGGTT 458
OY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198
||| ||| ||| ::::| ||||| ||||| ||||| ||||| |||||
DB 459 AAGTCAACAATGAGAGATATGTTGCTCTTGCACACTATGATCTCCAGGTAAATGTCATA 518
OY 199 GlyGlnSerProTyr 203
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 519 GGCCAAGTCCATAC 533

```

RESULT 11

AAV72998 standard; cDNA; 771 bp.

AAV72998;

25-FEB-1999 (first entry)

PR-1b protein encoding cDNA sequence.

Regulation: transcription; plant tissue; chimeric construction; PR.

KW pathogenesis-related protein; anti-pathogen; transgenic plant;

KW beta-1,3-glucanase activity; pest resistance; ss.

OS Nicotiana sp.

US5847258-A.

08-DEC-1998.

31-MAY-1995; 95US-0457364.

31-MAY-1995; 95US-0457364.

08-MAR-1988; 88US-0165667.

06-FEB-1989; 89US-0305566.

24-MAR-1989; 89US-0329018.

20-JUN-1989; 89US-0368672.

20-OCT-1989; 89US-0425504.

07-SEP-1990; 90US-0580431.

21-DEC-1990; 90US-0632441.

01-APR-1991; 91US-0678378.

27-SEP-1991; 91US-0768122.

06-NOV-1992; 92US-0848506.

06-NOV-1992; 92US-0973197.

06-APR-1993; 93US-0042847.

12-APR-1993; 93US-0045957.

16-JUL-1993; 93US-0093301.

13-JAN-1994; 94US-0181271.

(NOVS ) NOVARTIS FINANCE CORP.

Moyer MB, Payne GB, Ryals JA, Ward ER;

WPI: 1999-059180/05.

PT DNA encoding pathogenesis-related glucanase proteins - useful for  
PT producing transgenic plants with enhanced disease or pest resistance  
XX  
PS Example 44: Column 175-176; 169pp; English.

XX The present invention describes a DNA molecule encoding a  
CC pathogenesis-related (PR) protein having beta-1,3-glucanase activity  
CC selected from PR-2, PR-2', PR-2'', PR-N, PR-O and PR-O'. Also described  
CC are: (i) a chimeric gene comprising the above DNA molecule linked to a  
CC heterologous promoter; (ii) a vector containing the chimeric gene;  
CC (iii) a host cell containing the chimeric gene; (iv) a transgenic plant  
CC containing the chimeric gene; and (v) a seed from the transgenic plant.  
CC The DNA molecule is used to produce transgenic plants with enhanced  
CC disease or pest resistance. The present sequence represents a tobacco  
CC PR-1b protein encoding cDNA sequence from the present invention.

XX SQ Sequence 771 BP; 227 A; 148 C; 156 G; 240 T; 0 other;

Alignment Scores:

Pre-No.:	2,45e-13	Length:	771
Score:	304.00	Matches:	64
Percent Similarity:	56.55%	Conservative:	18
Best Local Similarity:	44.14%	Mismatches:	57
Query Match:	28.02%	Indels:	6
DB:	20	Gaps:	3

US-09-832-320-2 (1-203) x AAV72998 (1-771)

QY 61 SerGlySerAsnAlaThrAlaAspGluTyrLeuAlaProHisnglnAlaIaAla 80  
DB 111 TCCATGCCCCAAACTCTCAACAGACTTGTGGATGCCATACACAGCTCGTGAGAT 170  
QY 81 ValGlyValAlaProLeuArgTyrPasnAlaGlyLeuAlaSerAlaAlaAlaGlyThrVal 100  
DB 171 GTAGGCGTGAACACATTAACTTGGACACAGGGAGACGACCTATGCACAAATTATGTT 230  
QY 101 AlAGlnGlnArgArgGlnGlyGlyCysAlaPheAlaSpValGlyAlaSerProTyrGly 120  
DB 231 TCTCAATTGGCTGACAGCTGCACCTCGTACATTCTCAT-----GCCCAATGCGC 281  
QY 121 AlaAsn-----GlnGlyTyrPalaSerTyrArgAlaArgProAlaGluValAlaLeu 138  
DB 282 GAATACCTAGCTCAGGAGTGGCGATTATGACG--GCTGCTAAGCGCGTCGAGATG 338  
QY 139 TyrValAlaGlnGlyArgTyrTyrThrHisAlaAsnThrCysAlaGlyArgGln 158  
DB 339 TGGGTGATGAGAACAGACTATGACCTGAATCTGTGTGCACAGACAGGTG 398  
QY 159 CysGlyThrTyrThrGlnValValTyrPargAsnThrAlaGluValGlyCysAlaGlnAla 178  
DB 399 TGTGACACTATACATCAGGTGGTGGCGTAACTCGGTGCTGTGATGCTGCTAGGTT 458  
QY 179 SerCysAlaThrGlyAlaThrLeuThrLeuCysLeuTyrAsnProHisGlyAsnValGln 198  
DB 459 AAGTGCACAAATGAGAGATATGTCTCTTGTGCACATGATGATCCAGGTAATGTCATA 518  
QY 199 GlyGlnSerProTyr 203  
DB 519 GCCCAAGTCCATAC 533

RESULT 12

ID AAC47173 standard; DNA; 779 BP.

XX AAC47173;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 52827.

XX Hybridisation assay; genetic mapping; gene expression control;

KM protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 09-MAR-1999; 99US-0123180.

XX 23-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 29-MAR-1999; 99US-0126264.

XX 01-APR-1999; 99US-0126785.

XX 06-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0128714.

XX 19-APR-1999; 99US-0129845.

XX 21-APR-1999; 99US-0130077.

XX 23-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 20-MAY-1999; 99US-0134768.

XX 21-MAY-1999; 99US-0134941.

XX 21-MAY-1999; 99US-0135124.

XX 25-MAY-1999; 99US-0135353.

XX 27-MAY-1999; 99US-0135629.

XX 28-MAY-1999; 99US-0136392.

XX 01-JUN-1999; 99US-0136782.

XX 03-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 14-JUN-1999; 99US-0138847.

XX 16-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 17-JUN-1999; 99US-0139453.

XX 18-JUN-1999; 99US-0139492.

XX 18-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

XX 18-JUN-1999; 99US-0139456.

XX 18-JUN-1999; 99US-0139457.

XX 18-JUN-1999; 99US-0139458.

XX 18-JUN-1999; 99US-0139459.

XX 18-JUN-1999; 99US-0139460.

XX 18-JUN-1999; 99US-0139461.

XX 18-JUN-1999; 99US-0139462.

XX 18-JUN-1999; 99US-0139463.

XX 18-JUN-1999; 99US-0139750.

XX 21-JUN-1999; 99US-0139763.

XX 21-JUN-1999; 99US-0139817.

XX 22-JUN-1999; 99US-0139899.

XX 23-JUN-1999; 99US-0140353.

XX 23-JUN-1999; 99US-0140354.

XX 24-JUN-1999; 99US-0140695.

XX 28-JUN-1999; 99US-0140823.

[illegible]

PR	13-SEP-1999;	99US-0155758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154772.
PR	22-SEP-1999;	99US-0155133.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155650.
PR	28-SEP-1999;	99US-01556458.
PR	29-SEP-1999;	99US-01556596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158363.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159363.
PR	14-OCT-1999;	99US-0159363.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Alignment Scores:			
Pred. NO.:	7.21e-13	Length:	779
Score:	297.50	Matches:	61
Percent Similarity:	57.78%	Conservative:	17
Best Local Similarity:	45.19%	Mismatches:	54
Query Match:	27.42%	Indels:	3
DB:	21	Gaps:	2
US-09-832-320-2 (1-203) x AAC47173 (1-779)			
OY	69	GIUTYCTLEUALPROHISASNGINLAARGALALAVAGIYVALALAPROLEUALRGTPR	88
DB	155	GACITTTTGGCAGTTCCACACCGAGCAGACGCCAGGTTGGGTGGACCCCTTAAGATGC	214
OY	89	ASNALGIYLEUALISERIALAIALAGIYTHVALAIGLNGHARGARGINGIYLY	108
DB	215	GACGAGAGAGGTGGCTGTATTCGCCGTAACATGTCTAACCC-----CGTAAAGGTGAC	268
OY	109	CYSALAPHEALASPYAGIYALASERPROTYGIDYALASNGINGIYTPRALASERTYR	128
DB	269	TGCGCTATGCAACACTCGACGGCGGCC--TATGAGAGAGAACCTGCGTTGGAGCAGCGGT	325
OY	129	ARGHLAGRPROALAGIYVALVALALALEUTPRVALAGIUGIYARGTYRTYRTHNIS	148
DB	326	AGCATGACAGCGGTGACGAGCAGTTCAGCATGTGGGTGGACGAGCAATTTGACACGATTTAT	355
OY	149	ALASASNSHTHCYSALALAGIYARGINGIYSGIYTHTYTTHGLNVALYITPRAG	168
DB	386	GATTCACATACATGTGCATGGGCAACAGTGTGGCCACTATACCTACAGTGTGTGGAGA	445



```

OY 169 AsnThrAlaGluValGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThrLeu 188
DB 446 AACCTCGACAGAGTGGGATGTGCAGAAAGTGAGATGCACAACTGTTATCATCT 505
OY 189 CysLeuTyrAsnProHisGlyAsnValGlnGlyGlnSerProTyr 203
DB 506 TGCACACTACGACTCTCCGCGTAAGTGGTGGCGAGTGGCTTAC 550

RESULT 13
AA006182
ID AA006182 standard; DNA; 696 BP.
XX
AC AA006182;
XX
DT 31-JAN-1991 (first entry)
XX
DE PR-1c cDNA cloned into plasmid pBSprl-312.
XX
KW Transgenic plants; disease resistance; chimeric DNA;
  plant pathogenesis-related protein; ss.
XX
OS synthetic.
XX
FH Key Location/Qualifiers
FT mat_peptide 120..533
FT /tag= a
FT /product=PR-1c
XX
PN EP392225-A.
XX
PD 17-OCT-1990.
XX
PE 21-MAR-1990; 90BP-0105336.
XX
PR 20-OCT-1989; 89US-0425504.
PR 24-MAR-1989; 89US-0329018.
PR 20-JUN-1989; 89US-0368672.
XX
PA (CIBA ) CIBA GEIGY AG.
XX
PI Ryals JA, Alexander DC, Goodman RM, Meins F, Payne GB;
PI Stinson JR, Neuhaus J-M, Moyer MB;
XX
DR WPI: 1990-313983/42.
DR P-PSDB; AAR08224.
XX
PT Disease-resistant transgenic plants - obt. using encoding an
PT Inducible pathogenesis-related protein from infected plants.
XX
PS Example 15; page 24; 77pp; English.
XX
CC This chimeric DNA sequence comprises the PR-1c plant pathogenesis-
CC related protein (PRP)-coding sequence (II), and a promoter sequence
CC which enhances transcription of (II). This construct is used to
CC produce transgenic plant cells or tissues with the ability to re-
CC generate into plants which are disease resistant.
CC See also AA006179-81, AA006183-86, AA006199-006208 and AA006829.
XX
SQ Sequence 696 BP; 197 A; 140 C; 140 G; 219 T; 0 other;

Alignment Scores:
Pred. No.: 8,18e-13 Length: 696
Score: 296.00 Matches: 60
Percent Similarity: 55.88% Conservative: 16
Best Local Similarity: 44.12% Mismatches: 36
Query Match: 27.28% Gaps: 2
DB: 11

US-09-832-320-2 (1-203) x AA006182 (1-696)
OY 69 GltTyrLeuAlaProHisGlnAlaArgAlaValGlyValAlaProLeuArgTyr 88
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 135 GACTATTGGATGGCCATTAACACAGCTCGTCAGATGTAGGTGTAACCTTGCCTGG 194
OY 89 AsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAlaGlnGlnArgGlnGly 108
DB 195 GACGACGAGTGGAGCGCTATGCACAAATTTATCTCCCAATGGCTGCAGATTGTAA 254
OY 109 CysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTyrAlaSerTyr 128
DB 255 CTCGTACATCTGCAT-----GTCATACGCGGAGAAACCTATCTTGGGAAAGTGGC 305
OY 129 Arg---AlaArgProAlaGluValAlaLeuTyrValAlaGluGlyArgTyrTyrThr 147
DB 306 GATTCTTGGACGGCCGTAAAGCCGTCGAGATGTGGGTCAATGCAAAACGATTATGCC 365
OY 148 HisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyrThrGlnValTyrP 167
DB 366 CACGACTCAACACTTGTGCGCCAGACAGGTGTGGACACTATACCTCAGGTGGTGG 425
OY 168 ArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThrLeuThr 187
DB 426 CGTAACTCGCTTCGTGTGATGTGCTAGGGCTCAGTTCAGTTCATGCAATGAGAGATATGTGTC 485
OY 188 LeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGlnSerProTyr 203
DB 486 TCTTGCACTATGATCTCTCAGTAATGTATAGCCAAAGCCCATAC 533

RESULT 14
AAV62807
ID AAV62807 standard; cDNA; 696 BP.
XX
AC AAV62807;
XX
DT 05-MAR-1999 (first entry)
XX
DE Tobacco PR-1c gene clone.
XX
KW Chemically regulatable DNA promoter; expression control; pesticide;
  herbicide tolerance; pathogenesis related gene; PR gene; ss.
XX
OS Nicotiana acuminata.
XX
PN US5851766-A.
XX
PD 22-DEC-1998.
XX
PE 31-MAY-1995; 95US-0456262.
XX
PR 31-MAY-1995; 95US-0456262.
XX
PA (NOVUS ) NOVARTIS FINANCE CORP.
XX
PI Harms C, Ryals JA;
PI
XX
DR WPI: 1999-080396/07.
XX
PT Isolating chemically regulatable DNA sequences in plants - useful
PT for chemically controlling expression in transformed plants
XX
PS Example 44; Column 181-184; 175pp; English.
XX
CC This sequence represents a clone of the tobacco pathogenesis related (PR)
CC gene. This gene can be isolated using the method of the invention.
CC The method is for isolating a chemically regulatable DNA promoter
CC fragment from the 5' flanking region of a chemically regulatable gene in
CC a plant tissue. The method allows isolation of sequences which will be
CC useful for the controlled expression of genes, under the control of a
CC non-coding regulatable sequence. This is useful in plants with a
CC herbicide or pesticide detoxification mechanism under the control of a
CC chemical regulator, the regulator being applied before or with the
CC herbicide or pesticide to give optimal tolerance. The promoter fragment
CC is useful for controlling sequences which encode traits such as
CC height, shape, development, male or female sterility, and the ability
CC of the plant to withstand cold, heat, salt and drought. The chemical

```

CC induction of the promoter allows the regulation of production of  
CC compounds, e.g. flavours, fragrances, pigments, natural sweeteners,  
CC industrial feedstocks, antimicrobials and pharmaceuticals, by  
CC biosynthesis or metabolite conversion, whose biosynthesis is controlled  
CC by endogenous or foreign genes. The method allows control over the time  
CC and rate of gene expression either throughout the whole plant, or in  
CC localized tissues, to achieve e.g. fungal or insect resistance by for  
CC instance dusting the leaves with the chemical regulator. Controlling the  
CC developmental processes by the application of a regulating chemical in  
CC e.g. the commercial production of cultivated crops allows processes such  
CC as germination, flower formation and fruit ripening to be synchronised at  
CC a given time.

XX Sequence 696 BP; 197 A; 140 C; 140 G; 219 T; 0 other;

Alignment Scores:

Pred. No.:	8,18e-13	Length:	696
Score:	296.00	Matches:	60
Percent Similarity:	55.88%	Conservative:	16
Best Local Similarity:	44.12%	Mismatches:	56
Query Match:	27.28%	Indels:	4
DB:	20	Gaps:	2

US-09-832-320-2 (1-203) x AAV62807 (1-696)

OY 69 GUTYRLEUALAPROHISASNGIALAARGALALVALGLYVALAProleuAArgTrr 88  
DB 135 GACTATTGAGTGGCCATTAACACAGCTCGGACAGATGATGATGACCTTTGACCTGG 194  
OY 89 ASDAAGLYLEUALASERALALALAGLYTHRVAlALAGLNGlnARgAGlnGly 108  
DB 195 GACGACCGAGTACGCTATGCAACAAATTATGCTCCCAATTGGCTGACAGATTGTAC 254  
OY 109 CysAlaIpheAlaAspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTrrAlaSerTyr 128  
DB 255 CTCGACATTTCAT-----GGTCATACGGCGCAAAACCTGACCTTGGGAAAGTGGC 305  
OY 129 Arg---AlaArgProAlaGluValAlaAlaLeuTrrPValAlaGluGlyArgTyrTrrThr 147  
DB 306 GATTTCCTTGAGCGCCCTAAGGCCGTGAGATGTCGATGCAATGAGCAACAGTATTATCC 365  
OY 148 HsAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyrThrcInValValTrrP 167  
DB 366 CACGACTCAAACTTGTGTGCGCAAGACAGAGTGTGGACACTATACAGTGGTGG 425  
OY 168 ArgAsnThrAlaGluValGlyCysAlaAlaSerCysAlaThrGlyAlaThrLeuThr 187  
DB 426 CGTAACTCGGTTCGTGTGATGCTAGCGCTTCACTGTACAAATGAGAGATATATTGTC 485  
OY 188 LeuCysLeuTyrAsnProHisGlyAlaAsnValGlnGlyAlaSerProTyr 203  
DB 486 TCTTCCACACTATGATCCTCCAGGTAATGTTATATGCGCAAAAGCCCATAC 533

RESULT 15

AAV72999 standard; cDNA; 696 BP.

AAV72999;

25-FEB-1999 (first entry)

PR-1c protein encoding cDNA sequence.

Regulation; transcription; plant tissue; chimeric construction; PR;  
pathogenesis-related protein; anti-pathogenic; transgenic plant;  
beta-1,3-glucanase activity; pest resistance; ss.

Nicotiana sp.

US5847258-A.

08-DEC-1998.

PF 31-MAY-1995; 95US-0457364.  
XX 31-MAY-1995; 95US-0457364.  
PR 08-MAR-1988; 88US-0165667.  
PR 06-FEB-1989; 88US-0165667.  
PR 24-MAR-1989; 89US-0329018.  
PR 20-JUN-1989; 89US-0368672.  
PR 20-OCT-1989; 89US-0425504.  
PR 07-SEP-1990; 90US-0580431.  
PR 21-DEC-1990; 90US-0632441.  
PR 01-APR-1991; 91US-0678378.  
PR 27-SEP-1991; 91US-0768122.  
PR 06-MAR-1992; 92US-0848506.  
PR 06-NOV-1992; 92US-0973197.  
PR 06-APR-1993; 93US-0042847.  
PR 12-APR-1993; 93US-0045957.  
PR 16-JUL-1993; 93US-0093301.  
PR 13-JAN-1994; 94US-0181271.

PA (NOVS ) NOVARTIS FINANCE CORP.

PI Moyer MB, Payne GB, Ryals JA, Ward ER;

WPI; 1999-059180/05.

PT DNA encoding pathogenesis-related glucanase proteins - useful for  
producing transgenic plants with enhanced disease or pest resistance

PS Example 44; Column 177-178; 169pp; English.

XX The present invention describes a DNA molecule encoding a  
XX pathogenesis-related (PR) protein having beta-1,3-glucanase activity  
XX selected from PR-2, PR-2', PR-N, PR-O and PR-O'. Also described  
XX are: (i) a chimeric gene comprising the above DNA molecule linked to a  
XX heterologous promoter; (ii) a vector containing the chimeric gene;  
XX (iii) a host cell containing the chimeric gene; (iv) a transgenic plant  
XX containing the chimeric gene; and (v) a seed from the transgenic plant.  
XX The DNA molecule is used to produce transgenic plants with enhanced  
XX disease or pest resistance. The present sequence represents a tobacco  
XX PR-1c protein encoding cDNA sequence from the present invention.

SQ Sequence 696 BP; 197 A; 140 C; 140 G; 219 T; 0 other;

Alignment Scores:

Pred. No.:	8,18e-13	Length:	696
Score:	296.00	Matches:	60
Percent Similarity:	55.88%	Conservative:	16
Best Local Similarity:	44.12%	Mismatches:	56
Query Match:	27.28%	Indels:	4
DB:	20	Gaps:	2

us-09-832-320-2 (1-203) x AAV72999 (1-696)

OY 69 GUTYRLEUALAPROHISASNGIALAARGALALVALGLYVALAProleuAArgTrr 88  
DB 135 GACTATTGAGTGGCCATTAACACAGCTCGGACAGATGATGACCTTTGACCTGG 194  
OY 89 ASDAAGLYLEUALASERALALALAGLYTHRVAlALAGLNGlnARgAGlnGly 108  
DB 195 GACGACCGAGTACGCTATGCAACAAATTATGCTCCCAATTGGCTGACAGATTGTAC 254  
OY 109 CysAlaIpheAlaAspValGlyAlaSerProTyrGlyAlaAsnGlnGlyTrrAlaSerTyr 128  
DB 255 CTCGACATTTCAN-----GGTCATACGGCGCAAAACCTGACCTTGGGAAAGTGGC 305  
OY 129 Arg---AlaArgProAlaGluValAlaAlaLeuTrrPValAlaGluGlyArgTyrTrrThr 147  
DB 306 GATTTCCTTGAGCGCCCTAAGGCCGTGAGATGCTGATGATGACAAAGATATATTGTC 365  
OY 148 HsAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyrThrcInValValTrrP 167  
DB 366 CACGACTCAAACTTGTGTGCGCAAGACAGTGTGGACACTATACAGTGGTGG 425

Oy 168 ArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAlaThrGlyAlaThrLeuThr 187  
|||::: |||::: |||:::  
Db 426 CGTAACTCGGTTCGTGGATGTGCTAGGGTTCAGTGAACATGAGAGATATATGTC 485  
Oy 188 LeuCysLeuTyrAsnProHisGlyAsnValGlnGlyGlnSerProTyr 203  
|||::: |||::: |||:::  
Db 486 TCTTGCAACTATGATCCTCCAGGTAAAGTATAGGCAAAAGCCCATAC 533

Search completed: February 22, 2003, 03:07:51  
Job time : 235 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 22, 2003, 02:55:36 : Search time 1470 Seconds

(without alignments)  
2236.520 Million cell updates/sec

Title: US-09-832-320-2

Perfect score: 1085

Sequence: 1 MAHSSHHHLLLPAPMAFA.....ATITLCLYNPHGVGOSPY 203

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 6.0 , Ygapext 0.5  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPTO.spool/US0983320/runat.20022003.092737.14802/app.query.fasta.1.391  
-DB-RSP -OEMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -IOOPC=0 -IOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdd -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTWT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US0983320.0CGN.1.1.2874@runat.20022003.092737.14802 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -GAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	639	58.9	996	12	BG343299
2	568	52.4	362	13	BM101325
3	411.5	37.9	681	9	AU082529
4	402	37.1	666	13	BM259813
5	387.5	35.7	682	14	BQ411000
6	384	35.4	640	14	BQ010347
7	384	35.4	655	14	BQ015216
8	384	35.4	657	14	BQ011556
9	384	35.4	657	14	BQ013102
10	384	35.4	729	14	BQ986980
11	382	35.2	653	14	BQ013031
12	382	35.2	656	14	BQ013525
13	381.5	35.2	563	9	AT731816
14	380.5	35.1	637	12	BF268400
15	380	35.0	607	14	U47093
16	367.5	33.9	607	10	AM094514
17	366	33.7	653	14	BQ114041
18	364	33.5	639	14	BQ114394
19	363	33.5	551	9	AJ497996
20	358	33.0	500	14	BQ252852
21	355	32.7	489	14	BQ114395
22	349.5	32.2	621	14	BQ411001
23	349	32.2	706	13	B1310259
24	348.5	32.1	628	10	AV782649
25	348.5	32.1	673	10	AV822028
26	346	31.9	660	14	BM813422
27	342	31.5	427	12	BG451575
28	342	31.5	486	10	BE512404
29	340.5	31.4	533	12	BQ442275
30	337	31.1	654	10	BE600449
31	337	31.1	637	10	BE367243
32	337	31.1	759	11	AV106735
33	336.5	31.0	740	10	BE034124
34	333.5	30.7	554	10	AM037477
35	328	30.2	453	14	BQ114042
36	328	30.2	539	10	AM678619
37	328	30.2	588	10	BE597194
38	328	30.2	593	10	AM678759
39	328	30.2	611	10	BE367671
40	327	30.1	293	9	AU029886
41	325.5	30.0	578	13	B1124974
42	321.5	29.6	556	13	B1180121
43	320.5	29.5	752	17	BH449303
44	320	29.5	539	10	BE367566
45	320	29.5	554	10	AM746957

# ALIGNMENTS

RESULT 1  
BG343299  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

BG343299 996 bp mRNA linear EST 22-OCT-2001  
HYCDNA0008 (white to yellow anther) Hordeum vulgare pre-anthesis spike EST library  
HYCDNA0008 (white to yellow anther) Hordeum vulgare cDNA clone  
BG343299  
BG343299.1 GI:13155628  
EST.  
Hordeum vulgare.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
1 (bases 1 to 996)



```
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
```

(m.Fregene@cglar.org)  
Seq primer: Cassava custom 3' primer  
High quality sequence stop: 515.  
Location/Qualifiers

## FEATURES

1.666  
/organism="Manihot esculenta"  
/db\_xref="taxon:3983"  
/clone\_lib="Cassava EGC library1"  
/lab\_host="DH5alpha (E.coli)"  
/note="Vector: pYES2; Site 1: NotI; Site 2: EcoRI; The  
Cassava EGC library1 cDNA library was constructed using  
the directional cloning kit of Amersham Pharmacia and is  
directionally cloned into the EcoRI and NotI sites of the  
vector pYES2. This library was constructed by M. Fregene  
and R. Teranuchi."

BASE COUNT 172 a 157 c 158 g 179 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.28e-27 Length: 666  
Score: 402.00 Matches: 89  
Percent Similarity: 59.14% Conservatve: 21  
Best Local Similarity: 47.85% Mismatches: 62  
Query Match: 37.05% Indels: 14  
DB: 13 Gaps: 5

US-09-832-320-2 (1-203) x BM259813 (1-666)

```

OY 20 AlAcySleuLeuAlaThrlleuAlaLeuCySaAlaAlaProAlaProhrHlaSgLy 39
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 647 TCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG
OY 40 AlAaArgValleuMetProGlyGlyValaGlyAlaValaThrlYsaAlaGlnGlnGlyThr 59
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 599 TCT-----TCCACGGTCCACACAGCCACACAGCCG-----AACG 564
OY 60 GlySerGlySerAsnAlaThrAlaAspGlyThrLeuAlaProHisAsnGlnAlaArg 79
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 563 CCCACCGCTCTCCACACCGCCGCGCGCAATTCCTAGAACACACACACAGCAGCA 504
OY 80 AlAaGlyValAlaProLeuAlaThrlleuAlaLeuCySaAlaAlaProAlaProhrHlaSgLy 99
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 503 GCAGTGGGGGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 444
OY 100 ValAlaGlnGlnArgArgGlnGlyGlyCySaAlaAlaAspValGlyAlaSerProThr 119
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 443 GTCAGGTACCAAGAAAGAAATGGGTGTCAATTTGCCAATTTACACCTCTCTCTAC 384
OY 120 GlyAlaAsnGlnGlyThrAlaSerTyArgAla---ArgProAlaGluValAlaAlaLeu 138
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 383 GGTGGAACCAACCACTATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 324
OY 139 TrpValAlaGlnGlyArgTyThrHisAlaAsnAsnThrCySaAlaAlaGlyArgGln 158
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 323 TGGGTCTCAAGAAAGATTTACTACAACTCTATTAATCTCTGTGTCCCAATTCACACG 264
OY 159 CySgLyThrTyThrGlnValAlaTrpArgAsnThrAlaGluValGlyCySaAlaGlnAla 178
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 263 TGGCGGTCTTAACTCAGGTTGTGTGAAGAAGTCTCGAATGGGATCGGACACAGCT 204
OY 179 SerCySaAla---ThrGlyAlaThrLeuThrLeuCySleuTyAsnProHisGlyAsnVal 197
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 203 GCTTGTCTTAAGACGACGAGCTTAAGTGTGTCTTATTAATCCCTCGGGAATGT 144
OY 198 GlnGlyGlnSerProThr 203
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 143 GTTGGAGAGAGCCCATAT 126

```

RESULT 5 BQ411000 682 bp mRNA linear EST 22-MAY-2002  
LOCUS BQ411000 GA\_Ed0036A09f Gossypium arboreum 7-10 dpa fiber library Gossypium  
DEFINITION arboreum cDNA clone GA\_Ed0036A09f, mRNA sequence.

ACCESSION BQ411000  
VERSION BQ411000.1 GI:21098687  
KEYWORDS EST.  
SOURCE Gossypium arboreum.  
ORGANISM Gossypium arboreum.

REFERENCE 1 (bases 1 to 682)  
Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry  
D., Wood,T.C., Leslie,A. and Wilkins,T.A.  
An integrated analysis of the genetics, development, and evolution  
of the cotton fiber  
Unpublished (2000)  
CONTACT: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
1.682  
/organism="Gossypium arboreum"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_Ed0036A09f"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/tissue\_type="Fibers isolated from bolls harvested 7-10  
dpa"  
Total High Quality bases = 580  
Seg primer: TATACGACTCCTATAGCG  
High quality sequence start: 4  
High quality sequence stop: 643.  
Location/Qualifiers

BASE COUNT 215 a 159 c 140 g 167 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 5.15e-26 Length: 682  
Score: 387.50 Matches: 82  
Percent Similarity: 54.04% Conservatve: 25  
Best Local Similarity: 41.41% Mismatches: 64  
Query Match: 35.71% Indels: 27  
DB: 14 Gaps: 4

US-09-832-320-2 (1-203) x BQ411000 (1-682)

```

OY 14 ProAlaProMetAlaThrAlaCySleuLeuAlaThrlleuAlaLeu----- 30
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 16 CCAACACCCAAATGGCTTACCGCTGTGATTCATTTCTAGCTCTAGCCATTACACAC 75
OY 31 -----CySaAlaAlaProAlaPro---ThrHisGlyAlaArgValleuMetProGly 46
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 76 ACGGCCAAAGTCCACACACACACACACACACACACACACACATCAATGCTGCCA--- 132
OY 47 GlyAlaGlyAlaValaThrlYsaAlaGlnGlnGlyThrGlySerGlySerAsnAlaThr 66
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 133 -----CCAGCA 138
OY 67 AlaAspGlyThrLeuAlaProHisAsnGlnAlaArgAlaAlaValaGlyAlaProLeu 86
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 139 GCCAGACACTCTCTCCAGGCCACACCAAGCAAGAGCTGAGTGTGCTGCTCTCTCTCT 198
OY 87 ArgTrpAsnAlaGlyLeuAlaSerAlaAlaAlaGlyThrValAlaGlnGlnArgGln 106
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 199 AAATGGAGCCAAACACTAGCAATGCTTCAAGTCTCATCGCAAGGTACCAAGGACAAA 258
OY 107 GlyCySaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 126
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 259 ATGGGTGTCTGATTCGCAAACTGACAAACCAATTTAGCGAGCAACCAATTTAGGGGT 318

```



QY	127	SetTyrAlaA----	ArgProIacIuValAlaIeUtrPAlaIacIuGlyArgTyr	145
Db	319	AGCGAGCCGCCGTGACGCCGTCA	TGGCGGTGGAAACTTGGGTAAAGAGAAATTTT	378
QY	146	TyTThiAlaAsnaSnThrcCysAlaIaIacIyArgIncCysGlyThrTyThrcIuVal		165
Db	379	TATGATATATGCGACTATATCTGTGTCACCGCATCATATATATATGCTGTATATACACAGTT		438
QY	166	ValTTrArAsnThrcIacIuValAlcCysAlaIacIuAlaSerCysAlaThrGlyAlaThr		185
Db	439	GTTTGGAGAGATTTCTTCGATTGGGTGCTGCACAAACCTACATGCAAGATCAAGTACT		498
QY	186	LeuThrLeuCysLeuTyTranProHicIcysAlaValGInGlyInSerProTyr		203
Db	499	TTAACTATTTGTTTTATATATCTCTGCTGATATATATATAGTAGAAGAACATTC		552
RESULT 6				
LOCUS	BU010347		640 bp	MRNA linear EST 22-AUG-2002
DEFINITION	OGJ13C16.yg.ab1 OG_EFGHJ lettuce seriola Lactuca sativa cDNA clone			
ACCESSION	OGJ13C16.yg.ab1 OG_EFGHJ lettuce seriola Lactuca sativa cDNA clone			
VERSION	BU010347			
KEYWORDS	BU010347.1 GI:22444742			
SOURCE	EST.			
ORGANISM	Lactuca sativa.			
	Lactuca sativa.			
	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:			
	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:			
	Asteridae: euasterids II: Asterales: Asteraceae: Lactuceae;			
	Lactuca.			
REFERENCE	1 (bases 1 to 640)			
AUTHORS	Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,			
	Lin,H., van Damme,M., Lavelle,D., Chevallier,P., Ziegler,J., Ellison			
	,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,			
	Church,S., Jackson,L. and Bradford,K.			
TITLE	Lettuce and Sunflower ESTs from the Composite Genome Project			
	http://comgenomics.ucdavis.edu/			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Alexander Kozik [R.W.Michelmore]			
	Department of Vegetable Crops, R.W.Michelmore Lab			
	University of California at Davis (UCD)			
	Amundson Hall, UCD, Davis, CA 95616, USA			
	Tel: 1-(530)-742-1742			
	Fax: 1-(530)-752-9659			
	Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]			
	belongs to contig QG_CA.Contig198161, see http://cgpdb.ucdavis.edu/			
	for details.			
	Plate: QGJ13 row: C column: 16.			

FEATURES					
	source		Location/Qualifiers		
	1.	640			
	/organism=	" <i>Tactusa sativa</i> "			
	/cultivar=	" <i>L. serricola</i> "			
	/db_xref=	"taxon:4236"			
	/clone=	"FGJ13C16"			
	/clone_1lb=	"QG_EFGHJ lettuce serricola"			
	/lab_host=	"E.coli"			
	/note=	Vector: pBRCDNASIIAB: The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <a href="http://cgpdb.ucdavis.edu/TAG_rlib-QG_EFGHJ_lettuce_serricola">http://cgpdb.ucdavis.edu/TAG_rlib-QG_EFGHJ_lettuce_serricola</a> TAG_TISSUE-flowers_pre-fertilized TAG_SDO-ecttGAcGGg"			
BASE COUNT	167	a	178	g	162 t
ORIGIN	133	c			

Alignment Scores:

Pred. No.:	9,956-26	Length:	60
Score:	384.00	Matches:	84
Percent Similarity:	55.56%	Conservative:	21
Best Local Similarity:	44.44%	Mismatches:	60
Query Match:	35.39%	Indels:	24
DB:	15	Gaps:	4

US-09-832-320-2 (1-203) x BU010347 (1-640)

[illegible]

RESULT	7
Locus	B0015216
DEFINITION	655 bp mRNA Linear EST 22-AUG-2002 OCS9J011_Y9_ab1 OG_EFGH lettuce serriola Lactuca sativa cDNA clone
ACCESSION	OCS9J011_mRNA sequence.
VERSION	B0015216
KEYWORDS	B0015216.1 GI:22449611
SOURCE	EST.
ORGANISM	Lactuca sativa. <i>Lactuca sativa</i> . Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids II; Asterales; Asteraceae; Lactuceae; <i>Lactuca</i> .
REFERENCE	1 (bases 1 to 655) Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riseberg,L., Lin,H., Van Damme,W., Lavelle,D., Chevalier,P., Ziegler,J., Ellison, P., Kolkmann,J., Sibaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K. Lettuce and Sunflower ESTs from the Compositae Genome Project <a href="http://comphenomics.ucdavis.edu/">http://comphenomics.ucdavis.edu/</a> Unpublished (2002)
JOURNAL COMMENT	Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore lab

University of California at Davis (UCD)  
 Asmumson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu (michelmore@ucdavis.edu)  
 belongs to contig QG\_CA.Contig19161, see <http://cpgdb.ucdavis.edu/>  
 for details.  
 row: J column: 11.  
 plate: QGJ9

BASE COUNT ORIGIN	a	c	g	t
----------------------	---	---	---	---

US-09-832-320-2 (1-203) x BU015216 (1-655)

[illegible][illegible]

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
B0011556		657 bp	OCJ16123	OG-FRGMJ	lettuce	serriola	linear
B0011556			OGJ16123		mRNA sequence.	Lactuca	EST 22-AUG-2002
B0011556			B0011556			saliva	cdna clone
B0011556			B0011556.1	GI:22445951			
EST							
Lactuca						saliva	
Lactuca						saliva	
Lactuca						saliva	

**REFERENCE**  
1 (Pages 1 to 657)  
**AUTHORS**  
Kozik, A., Michelmore, R.W., Knapp, S., Marvienko, M., Rieseberg, L.,  
Lin, H., van Damme, M., Iavelle, D., Chevaller, P., Ziegler, J., Ellison,  
P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,  
Church, S., Jackson, L. and Bradford, K.  
**TITLE**  
Letuce and Sunflower ESTs from the Composite Genome Project  
**JOURNAL**  
<http://compagnomics.ucdavis.edu/>  
**COMMENT**  
Contact: Alexander Kozik [R.W.Michelmore]

FEATURES	Location/Qualifiers
source	1. .657

US-09-832-320-2 (1-203) \* BU011556 (1-657)

```

source
1. .657
/organism="Lactuca sativa"
/cultivar="IL serritola"
/db_xref="taxon:4236"
/clone="QGJ3124"
/clone.lib="QG-EPGHU lettuce serritola"
/lab_host="E.coli"
/note="Vector: pARCNDASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cspdb.ucdavis.edu/TAG\_LIB-QG-EPGHU lettuce serritola
TAG_MISUSE-flowers pre-fertilized
TAG_SEQ-GCTTGACGGG"
BASE COUNT      172 a      138 c      177 g      170 t
ORIGIN

Alignment Scores:
Pred. No.:      1,03e-25      Length:      657
Score:          384.00      Matches:      84
Percent Similarity: 55.56%      Conservative: 21
Best Local Similarity: 44.44%      Mismatches:   60
Query Match:     35.39%      Indels:       24
DB:              14      Gaps:         4

US-09-832-320-2 (1-203) x BU013102 (1-657)

OY      17  MetAlaThAlaGysLeuLeuAlaThrLeuAlaLeuAlaCysAlaAlaPro 36
      |||  |||  |||:::  |||:::  |||  |||  |||
Db      13  ATGACAAAGCTTCTTCCTTCACAGTACGTAGTATGGCCGCCACCGACCAAGCAGCG 72

OY      37  ThnHisGlyAlaArgValLeuMetProGlyGlyAlaGlyAlaValThrLysAlaGln 56
      |||  |||:::  |||  |||
Db      73  CCG-----GGGGGGGTGGGGCCGCCGAGGCTCAGGA----- 105

OY      57  GlyGlyThrGlySerGlySerSmaAlaThrAlaAspGluTyrLeuAlaProHisAsnGln 76
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      106 -----TATCTTGAAGCTCACACACAA 126

OY      77  AlaArgAlaValAlaGlyValAlaProLeuArgTrpAsnAlaGlyLeuAlaSerAlaAla 96
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      127  GCAGAGCCGACGTGGGTGGCTGCTCACTCAAGTGAAGACATCCAGCTGGTAAAGCAGC 186

OY      97  AlaGlyThrValAlaGlnGlnArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAla 116
      :::  |||  |||  |||  :::  |||  |||  |||  :::  |||
Db      187  AGCTTACTGGTGCATTCAGAGAGATCGACAAACATCCGATTGGCCACTTACAGAGC 246

OY      117  SerProTyrGlyValAsnGlnGlyTrpAlaSerTyrArg--AlaArgProAlaGluVal 135
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      247  GGGAGATGACGGCGGAACCAAGATGTGGGCCACCGGTGAGGTGTGACGCCGAGAGCGT 306

OY      136  ValAlaLeuTrpValAlaGluGlyArgTyrTrpHisAlaAsnAsnTrpCysAlaAla 155
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      307  GTGCACACGGGTGTCCGAGGAACGATTTTACCTTACGCAATATTTCAATGTGACACT 366

OY      156  GlyArgGlnCysGlyThrTyrTrpGlnValValTrpArgAsnThrAlaGluValGlyCys 175
      :::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      367  AGGCACCGGGGTGGGTGTACACGACGAGGGGTGTGGCGAAGCTGGCGAGAGCGGTGT 426

OY      176  AlaGlnAlaSerGlyAlaThrGly--AlaThrLeuTrpLeuGlySerLeuTyrAsnProHis 194
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      427  GCTTGGCTAGGTGTCTTAAGATCAGTCAACTTTCAGCATTTTCCTTTTAAATCCCTCC 486

OY      195  GlyAsnValGlnGlyInsSerProTyr 203
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      487  GGGATGTGTATCGAAGAACCCCTTAT 513

```





Db	185	ACGTTACGTGGGATTCACAGAGATGCAGAAAACGCCAGTTTGCAACTTAACGAC	244
Qy	117	SerProTyrGlyAlaAsnGlnGlyTrpAlaSerTyrArg---AlaArgProAlaGluVal	135
Db	245	GGGAGACGACGGCGGAAACGATGTGGGCACCGGTGGGTGTGACGCCGACGACGTG	304
Qy	136	ValAlaLeuTyrValAlaGluGlyArgTyrTyrThrHisAlaAsnThrCysAlaAla	155
Db	305	GTCGACACACGGGGTCCGGAGAACGCTATTACACTTACCCCATATATTATGTGACCT	364
Qy	156	GlyArgGlnCysGlyThrTyrThrGlnValValAlaTrpArgAsnThrAlaGluValGlyCys	175
Db	365	AGCCACCGGGGTGTGGGTGTACACGCCAGGTGGTGTGGCGAACTCGGCGGACGTGGGCTCT	424
Qy	176	AlaGlnAlaSerCysAlaThrGly---AlaThrLeuThrIleuCysLeuTyrAsnProHis	194
Db	425	GCTTGGCTGTAGGTGTCTAAAGATCAAGTCAACCTTAYGCGATTTGCTTATTATCTCCG	484
Qy	195	GlyAsnValGlnGlyIleSerProTyr	203
Db	485	GGGAAATGTGATCGAGAAAACCCCTTAT	511

RESULT 13			
AI731816			
LOCUS	AI731816	563 bp	mRNA
DEFINITION	BMHCH10971 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to STS1A PROTEIN PRECURSOR, mRNA sequence.		linear EST 11-JUN-1999

ACCESSION	AI731816
VERSION	AI731816.1
	GI:5050668

KEYWORDS	EST.
SOURCE	upland cotton.
ORGANISM	Gossypium hirsutum

REFERENCE  
1 (bases 1 to 563)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
1 (bases 1 to 563)

**AUTHORS** Brewitt,M., Matz,B.C., Davy,D.F. and Burr,B.  
**TITLE** ESTs from developing cotton fiber  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Ben Burr

Biology Department  
Brookhaven National Laboratory  
Upton, NY 11973, USA  
Tel: 516-344-3396  
Fax: 516-344-3407  
Email: [buriebniuxl.bnl.gov](mailto:buriebniuxl.bnl.gov)  
Seq primer: T3 Primer.

FEATURES	Location/Qualifiers
source	1. .563

BASE COUNT	ORIGIN
178 a	142 c 113 g 130 t

Alignment Scores:	
Pred. No.:	1,42e-25
Score:	381.50
Percent Similarity:	54.08%
Best Local Similarity:	41.03%
Query Match:	35.16%
DB:	9
Length:	563
Matches:	81
Conservative:	25
Mismatches:	63
Indels:	27
Gaps:	4

US-09-832-320-2 (1-203) x AI731816 (1-563)

QY 16 PrometalaThrAlaCysIeuLeuLeuA1aThrLeuLeuA1aLeu----- 30  
||| |||:: |||||:::  
Db 6 CCCAAATGGCCCTACGCCCTTGATTACATTTCTACTCTGACCATTTTACCAACACACGGCC 65

```

Qy      31 --CysAlaAlaPro---ThriSGlYAlaArgValLeuMetProGlyGlyAla 48
          |||||         |||
Ec GATTCTGGTCCAGCGACGATCGTCGCGGTGTTAATGCTGCTG-- 11

```

49 GlyAlaValThrLysAlaGlnGlnIgcIyThrgIySerGlySerAsnAlaThrAlaasp 68

Db 117 -----CCAGCAGCCACA 12

69 GATTGTLGLeuAlaProHsAsnGlnAlaArgAlaAlaValGlyValAlaProLeuAlaGlyP 88  
 :::::|||||  
 Db 129 GACTTCTCCAGCCCAACAACCAAGCAAGAGCTGCAGTAgGTGTCCGACACCTGTCAAAATGC 18

QY 89 AsnAlaClyLeuAlaSerAlaAlaAlaClyThrValAlaAlaGlnGlnArgGlnGly 10  
::: |||||:::||||::: ||||| :: |||

Db 189 AGCCACACACTAGCCAACTGCTTCAAGTCTCAATCCGCAAGTAACCAAGGACCAAAATGGGT 24

Qy 109 CysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsnGlnClylTTrpAlaSerTyr 122

Db 249 TGTGAGTTGCAACCTGACAAACCAAGTACGAGCAACACAGTTATGGGTAACGGA 30

Qy 129 ArgAla---ArgProAlaGluValValAlaLeuTrpValAlaGluLysArgTyrTyrThr 14

Db 309 GCCGCCGTACGCCCGTCATGGCGGTGGAACTTGGGTAAAGAGAAAGATTATATGAT 36

Qy 148 HisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThrTyrThrGlnValValTyr 16  
::: ||||| ||||| ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 369 TATGGGACTATACTTGTGCACCGAATCATAAATGTGGTGTATTACTCAAGTGTGG 42

Db 429 AAGAAATTCCTCGGATTGGGCTGCACACAGCTACATGCAGGATCAAGTTACTTAACT 48

Qy 188 LencysLeuTyrAsnProHisGlyAsnValGlnGlyGlnSerProTyr 203  
:::||||| ||||| ||::: |||||  
Db 489 APTGTGTTTATATCCCTCCTGGTAATTATATAGGTGAGAAACCATAC 536

RESULT 14  
BF268400

LOCUS	BF268400	637 bp	mRNA	linear	EST 07-MAR-2004
DEFINITION	GA_Eb0001L03f Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum cDNA clone GA_Eb0001L03f, mRNA sequence.				

ACCESSION	BF268400
VERSION	BF268400.1
KEYWORDS	GI:1199395 EST.

SOURCE	ORGANISM
Gossypium arboreum.	
Gossypium arboreum	
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophytes	

REFERENCE  
1 (bases 1 to 637)  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

**AUTHORS** Wang, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., H. D., Wood, T.C., Leslie, A. and Wilkins, T.A.

**TITLE** An integrated analysis of the genetics, development, and evolution of the

**JOURNAL**  
unpublished (2000)  
**CONTACT:** Wind PA

CONTACT: ming huang  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

100 Jordan Hall / Clemson / SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: [winia@clowson.edu](mailto:winia@clowson.edu)

Email: [twingec@emson.edu](mailto:twingec@emson.edu)  
Seq primer: TTAATGACTGACTATAGG  
High quality sequence stop: 629.

```

FEATURES      location/qualities
source        1. .637
              /organism="Gossypium arboreum"

```

```

/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Eb0001L03f"
/clone_1b="Gossypium arboreum 7-10 dpa fiber 1bira

```



Mon Feb 24 10:46:13 2003

us-09-832-320-2.p2n.rst

Page 12

Search completed: February 22, 2003, 04:05:00  
Job time : 1489 secs

20



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 22, 2003, 02:08:51 : Search time 1917 Seconds  
(without alignments)  
3081.831 Million cell updates/sec

Title: US-09-832-320-2

Perfect score: 1085  
Sequence: 1 MAHRSHHHLLLPAPMATA.....ATLTLCLYNNHGVGQSPY 203

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/gen2\_1/USPTO.spool/US09833320/runat.20022003.092737.14791/app.query.fasta.1.391  
-DB=genmb1 -GEMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NOR=ext -HEAPSIZEL=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09833320.0CGN\_1\_1.3745.0runat.20022003.092737.14791 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEJUNCTION -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=5 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Genmb1: \*  
1: gb\_ba: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pal: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*

29: em\_vl: \*  
30: em\_hlg\_hum: \*  
31: em\_hlg\_iny: \*  
32: em\_hlg\_other: \*  
33: em\_hlg\_mus: \*  
34: em\_hlg\_pln: \*  
35: em\_hlg\_rtd: \*  
36: em\_hlg\_mam: \*  
37: em\_hlg\_vrt: \*  
38: em\_sy: \*  
39: em\_higo\_hum: \*  
40: em\_higo\_mus: \*  
41: em\_higo\_other: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	699.5	64.5	144973	2	AC096689
2	699.5	64.5	150928	2	AC103550
3	365.5	33.7	1340	8	STST514
4	365.5	33.7	1340	8	STU17111
5	348.5	32.1	585	8	AY093248
6	348.5	32.1	624	8	AF386987
7	348.5	32.1	83511	8	AB013389
8	347.5	32.0	683	8	AY086114
9	331	30.5	716	8	ZM082200
10	325.5	30.0	707	8	AB015047
11	314.5	29.0	127053	8	HYPRIAR
12	313.5	28.9	155939	8	TIN24
13	313	28.8	155939	8	AP003853
14	312	28.8	784	8	ZMPRMS
15	310	28.6	110235	2	AP004057
16	309.5	28.5	110235	2	AP004057
17	309.5	28.5	153428	2	AP005535
18	308	28.4	1671	8	NTPRIB1
19	307.5	28.3	98124	8	ATP16L1
20	307.5	28.3	192861	8	ATCRIV80
21	306	28.2	864	8	AF306651
22	304	28.0	693	8	NTPRIB
23	304	28.0	735	6	I05690
24	304	28.0	735	6	I09280
25	304	28.0	735	6	NTPRIB
26	304	28.0	762	8	TOBPR1B1
27	304	28.0	768	8	NTPRIB
28	304	28.0	771	6	AR016785
29	304	28.0	771	6	AR020811
30	304	28.0	771	6	AR027134
31	304	28.0	771	6	AR038421
32	304	28.0	771	6	AR064563
33	304	28.0	771	6	AR067488
34	304	28.0	771	6	I38440
35	304	28.0	771	6	I56915
36	304	28.0	771	6	I59781
37	304	28.0	771	6	I75108
38	304	28.0	1149	8	NTPRIB1
39	304	28.0	1551	8	NTW381
40	303.5	28.0	2616	8	ATPRTAN
41	298.5	27.5	799	8	TAET348
42	297.5	27.4	492	8	AT129483
43	297.5	27.4	542	8	AF339699
44	297.5	27.4	687	8	HVPATRP1
45	297.5	27.4	747	8	AF324705

RESULT 1

## ALIGNMENTS

AC096689 144973 bp DNA linear HTG 09-JAN-2002  
LOCUS Oriza sativa chromosome 3 clone OSJNBa0027J18, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, 6 ordered pieces.  
ACCESSION AC096689.3 GI:18092993  
VERSION HTG: HTGS\_PHASE2.  
KEYWORDS Oriza sativa  
SOURCE Oriza sativa  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzaceae; Oriza.  
REFERENCE 1 (bases 1 to 144973)  
AUTHORS Buell, R., Hsiao, J., Zismann, V., Moffitt, K.M., Hill, J.,  
Gansberger, K., Burgess, S., Jarrahl, B., Shvartsbeyn, M., Brenner, M.,  
Ciecko, A., Pal, G., Vanaken, S., Hansen, C., Uterbach, T.,  
Feldblum, T., Khalak, H.G., Yuan, Q., Quackenbush, J., White, O.,  
Salzberg, S. and Fraser, C.  
TITLE Oriza sativa ssp. japonica cv. Nipponbare OSJNBa0027J18 BAC genomic  
sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 144973)  
AUTHORS Buell, R.  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
COMMENT On Jan 9, 2002 this sequence version replaced gi:17530731.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 29894: contig of 29894 bp in length  
\* 29895 29994: gap of unknown length  
\* 29995 70108: contig of 40114 bp in length  
\* 70109 70208: gap of unknown length  
\* 70209 79598: contig of 9390 bp in length  
\* 79599 79698: gap of unknown length  
\* 79699 106423: contig of 26725 bp in length  
\* 106424 106523: gap of unknown length  
\* 106524 139726: contig of 33103 bp in length  
\* 139727 144973: gap of unknown length  
\* 139727 144973: contig of 5247 bp in length.  
FEATURES  
source  
1. 144973  
Location/Qualifiers  
/organism="Oriza sativa"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="3"  
/clone="OSJNBa0027J18"  
BASE COUNT 38927 a 32302 c 32466 g 40767 t 511 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 3 92e-32 Length: 144973  
Score: 699.50 Matches: 145  
Percent Similarity: 72.528 Conservatve: 16  
Best Local Similarity: 65.328 Mismatches: 34  
Query Match: 64.47% Indels: 27  
DB: 2 Gaps: 7  
US-09-832-320-2 (1-203) x AC096689 (1-144973)  
OY 4 SerArgSerHisHisLeuLeu-----LeuLeuProAlaPro----- 16  
:::||||| ||| ||||| ||||| ::|||  
Db 135857 AATCGCGTTCACCTTCCTCACTGAGTGCACATCTACCTCAGATCTC 135916  
OY 17 -----MetAlaThrAlaCysLeuLeuAla 25

Db 135917 GCCATGCGCGCGCGGTGCTCGGTGCACGCGGCGCGCGGTCTCTCC 135976  
OY 26 ThrLeuAlaLeuCysAlaAlaProAlaProThrHisGlyAlaArgValLeuMetPro 45  
|||||::: ||||| ||||| |||||  
Db 135977 GCG---CTTGGCAGTGTGGCGCGCGCGCGCGCGGTGC---GGCGCGCGCGC---GTCTCT 136027  
OY 46 GlyGlyAlaGlyAlaValThrLysAlaGlnGlnGlyThrLysSerGlySerAsnAla 65  
|||||::: ||||| ||||| |||||  
Db 136028 CGTGCGCGCGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 136087  
OY 66 ThrAlaAspGlyUtyrLeuAlaProHisAsnGlnAlaArgAlaValGlyValAlaPro 85  
|||||::: ||||| ||||| |||||  
Db 136088 GCGCGCGAGCGTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 136147  
OY 86 LeuArgTyrPasnAlaGlyLeuAlaSerAlaAlaGlyThrValAlaGlnGlnArgArg 105  
|||||::: ||||| ||||| |||||  
Db 136148 CTGCGGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 136201  
OY 106 GlnGlyGly-----CysAlaPheAlaAspValGlyAlaSerProtyrGlyAla 121  
|||||::: ||||| ||||| |||||  
Db 136202 CAGGCGCGCGCGCGCGCGGTGCGCGGTGCGCGCGCGCGCGCGCGCGCG 136261  
OY 122 AsnGlnGlyTyrPalaSerTyrArgAlaArgProAlaGlyValAlaAlaLeuTyrValAla 141  
|||||::: ||||| ||||| |||||  
Db 136262 AACCGAGGAGTGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCG 136321  
OY 142 GlnGlyArgTyrTyrThrHisAlaAsnAspThrCysAlaAlaGlnGlnGlyThr 161  
|||||::: ||||| ||||| |||||  
Db 136322 CAGGAGAGGTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 136381  
OY 162 TyrThrGlnValAlaTyrPArgAsnThrAlaGlyValGlyCysAlaGlnAlaSerCysAla 181  
|||||::: ||||| ||||| |||||  
Db 136382 TACACCCAGAGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 136441  
OY 182 ThrGlyAlaThrLeuThrLeuCysAlaLeuTyrAsnProHisGlyValGlnGlyInsr 201  
Db 136442 ACCGCGCGCGCGCTACCATGTGCTCTACACCGCGCGCGCGCGCGCG 136501  
OY 202 ProTyr 203  
|||||  
Db 136502 CCTTAC 136507  
RESULT 2  
AC103550/c 150928 bp DNA linear HTG 04-JUN-2002  
LOCUS Oriza sativa chromosome 3 clone OSJNBa0079G12, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, 17 unordered pieces.  
ACCESSION AC103550  
VERSION AC103550.6 GI:21322019  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE Oriza sativa.  
ORGANISM Oriza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzaceae; Oriza.  
REFERENCE 1 (bases 1 to 150928)  
AUTHORS Buell, R., Hsiao, J., Zismann, V., Moffitt, K.M., Hill, J.,  
Gansberger, K., Burgess, S., Jarrahl, B., Shvartsbeyn, M., Brenner, M.,  
Ciecko, A., Pal, G., Vanaken, S., Hansen, C., Uterbach, T.,  
Feldblum, T., Khalak, H.G., Yuan, Q., Quackenbush, J., White, O.,  
Salzberg, S. and Fraser, C.  
TITLE Oriza sativa ssp. japonica cv. Nipponbare OSJNBa0079G12 BAC genomic  
sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 150928)  
AUTHORS Buell, R.  
TITLE Direct Submission  
JOURNAL Submitted (27-NOV-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
REFERENCE 3 (bases 1 to 150928)  
AUTHORS Buell, R.  
TITLE Direct Submission





DB	FEATURES	SOURCE
958	GGACAAACAAATTCGACGCTTGCATTAATTTAAAGTAAATGCGCAAAATATGATGCGCAATCAATT	957
Oy	124 YTPRALASSTYARV---AlaArgProAlaGluValValAlaLeuThrPheAlaIleGlu143	
Db	958 ATGGCGTAGAGGACGGTGTGACGCCACGAAATGGCTGTCAATCTTGGGGTCTGTAGAA1017	
Oy	143 YARGTYYTThHisAlaAsnAsnThrCysAlaAlaGluYARGLncCysGluThyTrpTh163	
Db	1018 GAAATTTAAATACATAAGAAATTAATTCATGACAGGGGATGATCAATAGTGGAGCTTAAAT1077	
Oy	163 rGlnValValTyrPArgAsnThrAlaGluValGluCysAlaGlnAlaAlaSerCysAlaThrcl183	
Db	1078 CCAAAATGTTGGAGAAATCAATGAAATTTGGCTGTGGCCAGACGTCTTACGAAAG1137	
Oy	183 y---AlaThrLeuThrLeuCysLeuThrAsnProHisGlnValGluIleGlnIleSer202	
Db	1138 ACGTCACTCTTACTGATGATGTTCTTATTAATCCAGCTGGAATATGATTAATTTAGAGAAACCC1197	
Oy	202 cTyr203	
Db	1198 TTAAT1201	
RESULT 5		
LOCUS	AY093248	
DEFINITION	Arabidopsis thaliana 585 bp mRNA linear PLN 21-APR--2002	
ACCESSION	AY093248	
VERSION	AY093248.1	
KEYWORDS	GI:20259799	
SOURCE	FLI, CDNA.	
ORGANISM	Arabidopsis thaliana.	
REFERENCE	Arabidopsis thaliana.	
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.1 (bases 1 to 585)	
	Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Bam, J., Carninci, P., Chen, H., Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, Y., Liu, S. X., Natusaka, M., Pham, P. K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Hayashizaki, K., Ecker, J., Theologis, A. and Davis, R. W.	
	Direct Submission	
	Submitted (26-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	
COMMENT	e-mail for correspondence: ardb@sequence.stanford.edu	
	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN Arabidopsis Full-length cDNA.) : Seki, M., Natusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
	The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick, A., Nguyen, M., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Bam, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, Y., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.	
	Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as pIs.	
	Location/Qualifiers	
	1..585	
	/organism="Arabidopsis thaliana"	
	/db_xref="taxon:3702"	
	/chromosome="5"	
	/clone="FJ15768"	
	/note="This clone is in pENTR/SD-dTropo. This is a cloned PCR product using RIKEN clone RAFL04-20-B17 (AF306987) as a template"	

gene	ecotype: Columbia"
1..585	/gene="At5g66590"
/note="X1P13.27"	
1..588	/gene="At5g66590"
/codon_start=1	
/product="unknown protein"	
/protein_id="AA13247.1"	
/db_xref="GI:20259800"	
/translation="MALTHIIFALLVTSKRAISPAALRKPROIVTSPPPTISA	
AKAKTADHNKARAAVGVPPVLSOTLEAAASLAAYONOKKCEFAALNGKXGANO	
IMAKTIVTPSLAVETWVKEKPYVAKKSTGTCAANTGTCYVQYVWRNRSLGCAQAT	
CRKESTVLTFTFVPRGNVIGOKPY"	
BASE COUNT	157 a 150 c 137 g 141 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.1e-13
Score:	348.50
Percent Similarity:	60.14%
Best Local Similarity:	48.25%
Query Match:	32.12%
DB:	8
	Gaps: 2
US-09-832-320-2 (1-203) x AY093248 (1-585)	
Oy 64 AaAAlaThrAlaAspSIuTYrLeuAlaProHIsaSnGlnAlaArGaAlaValGlyVal 83	
Db 127 TCTCCAGCAGCCTTAAAGCTTTCACCGATGCCGCAACAAGGCCAAGCCATGTTGGTCTT 186	
Oy 84 AlAProLeuArGrTYrPaAsnLaGlyLeuAlaSerAlaAlaIaGlyThrValAlaGlnGln 103	
Db 187 CCACCACTAGTTTGGAGCCAGACGTTGGAAAGTGTGGAGATGCGTGTGCTTACCG 246	
Oy 104 ArGrArGlnIaGlyValGlyAsAlaPheAlaAspValGlyAlaSerProTYrGlnAlaAsnGln 123	
Db 247 AGGAACCAAAAGAAAGATGTGATGTCGCGAGTCTAAACCTGGAAAATTCGGCGGAACCG 306	
Oy 124 GlyTPrAlaSer-----TYrArGrAlaArGrProAlaGlnValAlaAlaLeuTPrValAla 141	
Db 307 CTTTGGGCTTAAGGCTTACTAGCCCTGTGACACCGTCTCTGCTGGAGACGTTCGGTGAAG 366	
Oy 142 GlnGlyArGrTYrTYrThrHisAlaAsnAsnPrCysAlaAlaGlyArGlnCysGlyThr 161	
Db 367 GAGAAACCTTTCACATATATAAGTCAGACAGCATGTGTGTGGAAACCAACAGTCGGGATT 426	
Oy 162 TYrThrGlnValValITrPaGrAsnThrAlaGlnValGlyCysAlaGlnAlaSerCysAla 181	
Db 427 TATAAACAATGCTGTGGGAAACTCTAAAGAGCTCGGGGTGTGTCACAGCACGTGTACG 486	
Oy 182 ThrGlyAlaThr--LeuThrLeuCysLeuTYrAsnProHIsaGlnAsnValGlnGln 200	
Db 487 AAAGAGTCAACGGTGTWGCACATTTGTTTAAACATCTCCTGGAAATGTAAATWGGCAA 546	
Oy 201 SerProTYr 203	
Db 547 AAGCTTAC 555	
RESULT 6	
LOCUS AF386987 624 bp mRNA linear PLN 14-JUN-2001	
DEFINITION Arabidopsis thaliana Unknown protein (K1P13.27) mRNA, complete cds.	
ACCESSION AF386987.1 GI:14423499	
VERSION AF386987.1	
KEYWORDS F1.CDNA.	
SOURCE Arabidopsis thaliana.	
ORGANISM Arabidopsis thaliana.	
REFERENCE 1 (bases 1 to 624)	
AUTHORS Lam, B., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M.,	

Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shin, P., Yamada, K., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.  
 Direct Submission  
 Submitted (30-MAY-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
 The Salk, Stanford, FGENC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Lam, B., Karlin-Neumann, G., Nguyen, M., Southwick, A., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Pham, P.K., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shin, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.  
 Lam, B., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

# FEATURES

source  
 1. 624  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 /chromosome="V"  
 /clone="RAFLO4-20-B17, R15768"  
 /note="This clone is in plasmid"  
 ecotype: Columbia"  
 1. 624  
 /gene="KIF13.27"  
 10. 567  
 /gene="KIF13.27"  
 /product="unknown protein"  
 /protein\_id="AAK62432.1"  
 /db\_xref="GI:14423500"  
 /translation="MALTRHITFVALVIVSYKATSPAAKIKPKQIVSTSPPTTISA  
 AAATFDHAKKAMGVPRIVSQTLPAASRIARQKNCERASLNGRTGANO  
 LMAKGLVAVTSLAVETWKEKPEFYNGSDCANHTCGYKOVWMSKELCAQAT  
 CTKESTVLTICFNPNGNYIGQKPY"  
 BASE COUNT 175 a 152 c 143 g 154 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3 3e-13 length: 624  
 Score: 348.50 Matches: 69  
 Percent Similarity: 60.148 Conservative: 17  
 Best Local Similarity: 48.258 Mismatches: 54  
 Query Match: 32.128 Indels: 3  
 DB: 8 Gaps: 2

US-09-832-320-2 (1-203) x AF366987 (1-624)

64 AsnAaTrrAlaAaPrGluTyrLeuAlaProHisAsnGlnAlaAaGlaAlaValAlaGln 83  
 136 TCTGACGACGTAAAGCTTTACCGATGCGCACAAAGCCAGAGCCATGGTGGTGT 195  
 84 AlAPrLeuAaTrrPAsnAlaGlyLeuAlaSerAlaAlaGlyThrValAlaGln 103  
 196 CCACCATAGTTGGACGACGAGCTGGAGAGTGGTGGCTGGCTTACCAAG 255  
 104 ArgAaGlnGlyGlyCysAlaAlaAlaAaPValGlyAlaSerProTyrGlyAlaAsnGln 123  
 256 AGGAACCAAGAAAGTGGAGTGGAGTCTTAACCCCTGGAAATACGGCCGACCAAG 315  
 124 GlyTrrAlaSer-----TyrArgAlaAaTrrProAlaGluValAlaAlaLeuTrrPValAla 141

Db 316 CTTGGGCGTAAAGCGCTTAGTACCGGTGACACCGTCTTGGCTGGAGACTTGGGTAAAG 375  
 142 GluGlyAaTrrTrrAlaAlaAsnAlaSerAlaAlaGlyArgGlnGlyThr 161  
 376 GAGAAACCTTTCACATTTATAGTACAGACAGTGTGGTGGACACACAGTGGGGGT 435  
 162 TyrThrGlnValAlaTrrPAsnAlaGluValAlaGlyCysAlaGlnAlaSerCysAla 181  
 436 TATTAACAAAGCTCTGGAGAAACCTTAAGAGCTGGGTGTCTCAACACGACGTAGC 495  
 182 ThrGlyAlaThr--LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGly 200  
 496 AAAGTCAACAGGCTGACCATTTCTTTTACAAATCTCGGAAATGTATATGGCCAA 555  
 201 SerProTyr 203  
 556 AAGCCTTAC 564

RESULT 7  
 AB013389  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Arabidopsis thaliana (strain:Columbia) DNA, clone:lib:Mitsui TAC  
 clone:KIF13.  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eumossids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (sites)  
 Kotani, H., Nakamura, Y., Sato, S., Asamizu, E., Kaneko, T., Miyajima, N.  
 and Tabata, S.  
 Structural analysis of Arabidopsis thaliana chromosome 5. VI.  
 Sequence features of the regions of 1,367,185 bp covered by 19  
 physically assigned P1 and TAC clones  
 DNA Res. 5 (3), 203-216 (1998)  
 2 (bases 1 to 83511)  
 Nakamura, Y.  
 Direct Submission  
 Submitted (06-MAY-1998) Yasuoka, Nakamura, Kazusa DNA Research  
 Institute, Department of Plant Gene Research, 1532-3, Yana,  
 Kisarazu, Chiba 252-0812, Japan (E-mail: ynakamu@kazusa.or.jp,  
 Tel: 81-438-52-3935, Fax: 81-438-52-3934)  
 Address for correspondence: kaos@kazusa.or.jp  
 For the latest information on annotation of this clone, please see  
 http://www.kazusa.or.jp/kaos/cgi-bin/seq\_graph.cgi?c=KIF13

Genes with similarity to proteins in the databases are described in  
 'product' or 'note' qualifiers. Genes that have no significant  
 protein similarity are described as 'unknown protein'.  
 The software programs used to predict genes include: Graal  
 (Informatics Group, Oak Ridge National Laboratory,  
 http://compbio.ornl.gov/grail-1.3/),  
 GENSCAN (Chris Burge, MIT, http://ccr-081.mit.edu/GENSCAN.html),  
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
 SplicePredictor (Volker Brendel, Stanford University,  
 http://greenli.zool.iastate.edu/cgi-bin/sp.cgi).  
 Genes encoding tRNAs are predicted by tRNAcan-se  
 (Sean Eddy, Washington University School of Medicine, St. Louis,  
 http://genome.wustl.edu/eddy/tRNAcan-se/).  
 This sequence may not be the entire insert of this clone. It may be  
 shorter because we remove overlaps between neighboring submissions.  
 The 5' clone is K1120 and the 3' clone is MSN2.  
 location/Qualifiers

FEATURES  
 source  
 1. 83511  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="5"



/note="gene\_id:K1F13.12"  
/codon\_start=1

## Alignment Scores:

Pred. No.: 3,77e-11 Length: 83511  
Score: 348.50 Matches: 69  
Percent Similarity: 60.14% Conservative: 17  
Best Local Similarity: 48.25% Mismatches: 54  
Query Match: 32.12% Indels: 3  
DB: 8 Gaps: 2

US-09-832-320-2 (1-203) x AB013389 (1-83511)

QY 64 AsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaValGlyVal 83  
DB 63297 TCTGCAGCAGCTAAAGCTTCCACCGATGCCACACAGGCGAGCCATGTGTGTT 63356  
QY 84 AlaProLeuArgTyrPAsnAlaGlyLeuAlaSerAlaAlaGlyThrValAlaGln 103  
DB 63357 CCACCACTAGTTTGGAGCCAGACCTTGGAAGCTGCGAGTGGCTGCTTACCA 63416  
QY 104 ArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsn 123  
DB 63417 AGCAACCAAAAGAGTGTGATGCTCGCAGTCTAACCTGGAAATACGCGCAACAG 63476  
QY 124 GlyTPAlaSer-----TyrArgAlaArgProAlaGluValAlaLeuTyrValAla 141  
DB 63477 CTTTGGCTAAAGGCTTAAAGCCGTGACACCGTCTTCTGCTGCTGAGACTGGGTAA 63536  
QY 142 GluGlyArgTyrTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThr 161  
DB 63537 GAGAAACCTTCTCAATTAATTAAGTACAGACAGCTGCTGCGAACAACAGCGGGGT 63596  
QY 162 TyrThrGlnValValTyrPArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAla 181  
DB 63597 TATTAACCAAGTGTCTGAGAACTTAAAGAGCTGCGGCTGCTAACCCAGCTGTACG 63656  
QY 182 ThrGlyAlaThr---LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGly 200  
DB 63657 AAAGAGTCACGCTGTACCATTTGTTTACATCCCTCTGGAATGTAATGTGCCAA 63716  
QY 201 SerProTyr 203  
DB 63717 AAGCCTTAC 63725

## RESULT 8

AY086114

LOCUS AY086114 683 bp mRNA linear PLN 25-JUN-2002  
DEFINITION Arabidopsis thaliana clone 2152 mRNA, complete sequence.  
ACCESSION AY086114  
VERSION AY086114.1 GI:21404824  
KEYWORDS FLI\_CDNA.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 683)  
Haas,B.J., Volfovsky,N., Town,C.D., Troupkan,M., Alexandrov,N.,  
Feidmann,K.A., Flavell,R.B., White,O. and Szizberg,S.L.  
Full-length messenger RNA sequences greatly improve genome  
annotation

REFERENCE  
AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made available to RIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants; including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

## FEATURES

## source

## CDS

1..683  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
75..632  
/product="stsl4"  
/codon\_start=1  
/protein\_id="AA063321.1"  
/db\_xref="GI:21554246"  
/translation="MALTHITVALVTSVAISPAKLRKQIVSTSPPTTISA  
AAKFTDAHNRKAMVGPPLVWSQTEAAARLRVPRNRKCFASLNPGRKGAAG  
LMAKGLAATVPSLAVETWKEKPFYKEDTCAANHTGVYKQVWNRKSELCAQAT  
CTKESVTLTLCFNPONIGQKPY"  
BASE COUNT 197 a 163 c 148 g 175 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4.14e-13 Length: 683  
Score: 347.50 Matches: 68  
Percent Similarity: 60.14% Conservative: 18  
Best Local Similarity: 47.55% Mismatches: 54  
Query Match: 32.03% Indels: 3  
DB: 8 Gaps: 2

US-09-832-320-2 (1-203) x AY086114 (1-683)

QY 64 AsnAlaThrAlaAspGluTyrLeuAlaProHisAsnGlnAlaArgAlaValGlyVal 83  
DB 201 TCTGCAGCAGCTAAAGCTTCCACCGATGCCACAAAGGCGAGCCATGTGTGTT 260  
QY 84 AlaProLeuArgTyrPAsnAlaGlyLeuAlaSerAlaAlaGlyThrValAlaGln 103  
DB 261 CCACCACTAGTTTGGAGCCAGACAGCTGCGAGTGGCTGGCTGTATCCAG 320  
QY 104 ArgArgGlnGlyGlyCysAlaPheAlaAspValGlyAlaSerProTyrGlyAlaAsn 123  
DB 321 AGCAACCAAAAGAGTGTGATGCTCGCAGTCTAACCTGGAAATACGCGCAACAG 380  
QY 124 GlyTPAlaSer-----TyrArgAlaArgProAlaGluValAlaLeuTyrValAla 141  
DB 381 CTTTGGCTAAAGGCTTAAAGCCGTGACACCGTCTTCTGCTGAGACTGGGTAAAG 440  
QY 142 GluGlyArgTyrTyrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysGlyThr 161  
DB 441 GAGAAACCTTCTCAATTAATTAAGTACAGACAGCTGCTGCAACACAGCTGCGGGTT 500  
QY 162 TyrThrGlnValValTyrPArgAsnThrAlaGluValGlyCysAlaGlnAlaSerCysAla 181  
DB 501 TATTAACCAAGTGTCTGAGAACTTAAAGAGCTGCGGCTGCTCAAGCAGCTGTACG 560  
QY 182 ThrGlyAlaThr---LeuThrLeuCysLeuTyrAsnProHisGlyAsnValGlnGly 200  
DB 561 AAAGAGTCACGCTGTACCATTTGTTTACATCCCTCTGGAATGTAATGTGCCAA 620  
QY 201 SerProTyr 203







```

Db 241 ACCGAG---AGGATCGGCGACTGCAGCTCCAGACTCCGGC---GGGCGCTACGGGAG 294
OY 122 snngllytrpalaSerTYrArgAla-----ArgProAlaGluValAlaLeuTrpV 140
Db 295 ACATCTTCTGGGCTGCGCGCGGCGACTGGAGCGCGGACCGCGGTGAAGCTGTGGG 354
OY 140 alalagluglyArgTYrTYrThrHisAlaAsnAsnThrCysAlaAlaGlyArgGlnCysG 160
Db 355 TGGACGAGAGAGAGAGACTACGACTACGAGGTCACACACTGTGCGAGGGGAGGTGTGCG 414
OY 160 lythrTYrthrGlnValValTYrPArgAsnThrAlaGluValGlyCysAlaGlnAlaSerC 180
Db 415 GGCACCTACGACGAGTGTGGTGGCGGCTGCACCGACATCGGCTGCGCTGCTGCT 474
OY 180 ys---AlaTYrGlyAlaThrLeuThrLeuCysLeuTYrAsnProHisGlyAsnValGlnG 199
Db 475 GCACACAGACAGCGGCGGCTTCATCATCAGCTGCAGACTACGAGCCCGCGGAGATTTGTT 534
OY 199 lyGlnSerProTYr 203
Db 535 GACAGAAACCATAC 548

RESULT 12
TIN24 127053 bp DNA linear PLN 21-JUN-1999
LOCUS Arabidopsis thaliana BAC TIN24.
DEFINITION AF149413
ACCESSION AF149413.1 GI:4809270
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 127053)
Washington University Genome Sequencing Center.
The A. thaliana Genome Sequencing Project.
Unpublished (1997)
2 (bases 1 to 127053)
Murray, J., Langston, Y., Clarke, K. and Drone, K.
The sequence of A. thaliana TIN24
Unpublished (1999)
3 (bases 1 to 127053)
Waterston, R.
Direct Submission
Submitted (10-MAY-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
4 (bases 1 to 127053)
Waterston, R.
Direct Submission
Submitted (21-JUN-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: twatson@watson.wustl.edu

```

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems,

such as compressions and repeats; all regions were covered by sequence from more than one subclone

#### NEIGHBORING COSMID INFORMATION:

The 3' clone is F18A17. Actual start of this clone is at base position 1 of TIN24; actual end is at 127053 of TIN24

The 3' neighbor F18A17 was sequenced at the Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, ACCESSION # AC005405.

#### NOTES:

Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green and L. Hillier, ms in preparation).

#### FEATURES

##### source

##### gene

##### CDS

```

1..127053
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="V"
/map="unknown"
/clone="TIN24"
6032..8670
/gene="TIN24.15"
join(6032..6206,6502..6814,6901..7192,7282..7383,
7450..7878,7954..8670)
/gene="TIN24.15"
/note="contains similarity to protein kinase domains; Pfam
PF00069, score=129.3, E=7e-35, N=1"
/codon_start=1
/evidence-not_experimental
/protein_id="AAD40140.1"
/db_xref="GI:5107827"
/translation="WASPYSSDDIHSPVNSTVVAIDKEKSHSYAVRMADVHLNMIH
PVMILVHVRKNMNSDVAKTILDVYNNLVNVLVAGSTKTRFARSMFSKPREVQ
SISIMSTDFCSYVYISKGGVSSRPQKPIINTLAPRPSGFLIQLSDSDODL
IRVORSARNKPNETTYPHNRAAPNTQKPKPIINGSDPNNFGNOAFOGPILOS
SFSDSDGDFGVMGSDLSQNSMDYFHGASSSESIPOSTKDIIEARRLKLEIKO
TMDVTSACKEALFAKKRANELNQMKEEARKEKRAKLESEALAVAEIERAKCOTV
EAAEKAOMAELEGORRKOEMKAVSEKDRDAVSAALAHNRYRYKYSIEPEATE
RANRRKIGRGSGYVNGTIDHPVAIKYLRPAAGKROQOQEVPLCSTRHHMY
LLGACPEYCCIVFEFENSGLEDRLERTGNSPESWRKRFEIAELINAIASELQAK
PEPLVHRODLKPAINTILDKNVSKISDVGLAKLVPASIASVQFPMTSAGTFCYDIP
EYQDGMULTKSDVSGILLLQILITGRPPMGLAHQVSAISKGFKEMLDVPYDWP
VQDAQSFATLAKCAELKRDRLDGEVVPHLIRUKNGNDERTNEMI"
/gene="TIN24.16"
complement(join(14495..14593,14673..14816,14910..14979,
15134..15240,15324..15335))
/gene="TIN24.16"
/note="contains similarity to a Saccharomyces cerevisiae
hypothetical protein (SW:P47044) and a Bacillus subtilis
hypothetical protein (GB:U15180)"
/codon_start=1
/evidence-not_experimental
/protein_id="AAD40141.1"
/db_xref="GI:5107828"
/translation="WHEHISGEVGEVRIVSDMHRKATMAQEGAFATLIGERYET
MELEEMITMAGLGIHKKTGVLINVDGYVNNLAAPFDVGEGLFKQACNIVSAPS
AELEMKHELYTPSHKTYASHQSKVPELDYPLNENKPO"
17527..18021
/gene="TIN24.14"
17527..18021
/gene="TIN24.14"
/note="contains similarity to SCP-like extracellular
proteins; Pfam PF00188, score=196.7, E=3.7e-55, N=1"
/codon_start=1
/evidence-not_experimental
/protein_id="AAD40121.1"

```









Mon Feb 24 10:46:11 2003

us-09-832-320-2.p2n.rge

Page 16

Search completed: February 22, 2003, 03:42:25  
Job time : 2068 secs

-----



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 03:08:01 : Search time 233.123 Seconds  
(without alignments)  
8674.789 Million cell updates/sec

Title: US-09-832-320-1

Perfect score: 898  
Sequence: 1 ctccgacgcactgcacgctc.....aaaaaaaaaaaaaaaa 898

Scoring table: OLIGO.MDC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 25

Total number of hits satisfying chosen parameters: 30604

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq.101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	898	100.0	898	24	ABA96417
2	612	68.2	612	24	ABA96418
3	43	4.8	274	22	AAH35782
4	43	4.8	293	22	AAH71507
5	43	4.8	447	24	ABL93994
6	43	4.8	481	23	ABV58614
7	43	4.8	550	23	ABV57072
8	43	4.8	551	23	ABV56917
9	43	4.8	1077	22	AA527462

10	43	4.8	1091	22	AA099723	Maize ZmGnSN1-1 g1
11	43	4.8	1095	21	AA077435	Human ORFX ORF290
12	43	4.8	1240	24	ABL01589	Human secreted pro
13	43	4.8	1483	22	AA504152	Human ras converti
14	43	4.8	1549	20	AA024921	Human ras carboxy-
15	43	4.8	1821	19	AAV59590	Human secreted pro
16	43	4.8	1835	20	AA027245	Human CLAR1 coding
17	42	4.7	111	21	AA098625	Human colon cancer
18	42	4.7	190	23	ABV20053	Human prostate exp
19	42	4.7	198	23	ABV55195	Human prostate exp
20	42	4.7	198	23	ABV56628	Human prostate exp
21	42	4.7	272	23	ABV18469	Human prostate exp
22	42	4.7	319	23	ABV49812	Human prostate exp
23	42	4.7	342	23	ABV16183	Human prostate exp
24	42	4.7	395	22	AA187795	Human polynucleoti
25	42	4.7	407	23	ABV58185	Human prostate exp
26	42	4.7	424	23	ABV55712	Human prostate exp
27	42	4.7	443	8	AAW70773	Sequence encoding
28	42	4.7	456	22	AA183054	Human polynucleoti
29	42	4.7	464	23	ABV45981	Human prostate exp
30	42	4.7	465	23	ABV48253	Human prostate exp
31	42	4.7	495	22	ABAA3074	Human breast cell
32	42	4.7	495	22	ABAA3490	Human foetal liver
33	42	4.7	495	22	ABAA3257	Probe #1723 for ge
34	42	4.7	495	22	ABAA3257	Human brain expres
35	42	4.7	495	22	AAK01761	Human bone marrow
36	42	4.7	495	22	AAK27215	Probe #1729 for ge
37	42	4.7	495	22	AA131110	Probe #1796 used t
38	42	4.7	495	22	AA101729	Probe #1720 used t
39	42	4.7	495	22	AB501757	Human genome-deriv
40	42	4.7	595	24	AA501171	Fertilisation-Inde
41	42	4.7	707	22	AA525840	Human CDNA encodin
42	42	4.7	740	19	AAV23915	Plant CCR enzyme D
43	42	4.7	740	20	AA206881	Pine cinnamoyl-CoA
44	42	4.7	740	21	AA467915	Pinus radiata CCR
45	42	4.7	741	19	AAV23878	Plant CCR enzyme D

#### ALIGNMENTS

RESULT 1	ABA96417	standard; cDNA; 898 BP.
ID	ABA96417	
XX	ABA96417:	
AC		
XX		
DT	02-APR-2002	(first entry)
XX		
DE	Maize PR1-C10 encoding cDNA seq ID NO. 1.	
XX		
KW	Maize; pathogen-related; PR1-C10; plant; transgenic; gene; ss.	
XX		
OS	Zea mays.	
XX		
EH		
FT	Key	Location/Qualifiers
FT	CDS	63..674
FT		/*tag= a
FT		/Product= "PR1-C10"
XX		
XX	US2001049834-A1.	
XX		
PD	06-DEC-2001.	
XX		
FP	10-APR-2001; 2001US-0832320.	
XX		
PR	10-APR-2000; 2000US-195801P.	
XX		
PA	(CRAN/) CRANE E H.	
XX	(CRAN/) CRANE V C.	
XX	Human prostate exp	
XX	Human prostate exp	
XX	Crane EH, Crane VC;	
XX		

DR WPI: 2002-121407/16.  
 DR P-PSDB: AAM48742.  
 XX New nucleic acid encoding a pathogen-related protein isolated from  
 PT maize and designated PRI-C10, useful for transforming plants for  
 PT enhanced disease resistance  
 XX

PS Claim 1; Page 30-31; 34pp; English.

XX The invention relates to an isolated nucleic acid encoding a  
 CC pathogen-related protein PRI-C10. The nucleic acid is used to transform  
 CC plants for enhanced disease resistance.  
 XX

XX Sequence 898 BP; 161 A; 295 C; 309 G; 133 T; 0 other;

Query Match 100.0%; Score 898; DB 24; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-288;  
 Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGACGACACTGACGCTCATTCTACTGAGCCATTACTAGATCAGCACTCAGATCT 60  
 DB 1 CTCGACGACACTGACGCTCATTCTACTGAGCCATTACTAGATCAGCACTCAGATCT 60  
 QY 61 CAATGGGCGACTGGCGAGCGACACACCTCTCTGTCGCCGGCCAGGCGACAGG 120  
 DB 61 CAATGGGCGACTGGCGAGCGACACACCTCTCTGTCGCCGGCCAGGCGACAGG 120  
 QY 121 CGTGTTCCTCTGCGCCACCTCTCTGCGCTGCGCCGCCGCCGCCGCCGCCGCC 180  
 DB 121 CGTGTTCCTCTGCGCCACCTCTCTGCGCTGCGCCGCCGCCGCCGCCGCCGCC 180  
 QY 121 CGTGTTCCTCTGCGCCACCTCTCTGCGCTGCGCCGCCGCCGCCGCCGCCGCC 180  
 DB 121 CGTGTTCCTCTGCGCCACCTCTCTGCGCTGCGCCGCCGCCGCCGCCGCCGCC 180  
 QY 181 CGCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 DB 181 CGCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 QY 181 CGCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 DB 181 CGCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 QY 241 GCAGGCGGACGACGCGAGCGCGAGAGTACTGCGCGCGCAACACAGCGCGCGCG 300  
 DB 241 GCAGGCGGACGACGCGAGCGCGAGAGTACTGCGCGCGCAACACAGCGCGCGCG 300  
 QY 241 GCAGGCGGACGACGCGAGCGCGAGAGTACTGCGCGCGCAACACAGCGCGCGCG 300  
 DB 241 GCAGGCGGACGACGCGAGCGCGAGAGTACTGCGCGCGCAACACAGCGCGCGCG 300  
 QY 301 CGGTGGGCGGCG 360  
 DB 301 CGGTGGGCGGCG 360  
 QY 361 TGGCGGACGACG 420  
 DB 361 TGGCGGACGACG 420  
 QY 361 TGGCGGACGACG 420  
 DB 361 TGGCGGACGACG 420  
 QY 421 GCGGCAACGAGGGGTGGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 DB 421 GCGGCAACGAGGGGTGGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 QY 421 GCGGCAACGAGGGGTGGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 DB 421 GCGGCAACGAGGGGTGGGCGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 QY 481 TGGCGGAGGGGCGGTACTACACCGCGCAACACAGCGCGCGCGCGCGCGCGCGCG 540  
 DB 481 TGGCGGAGGGGCGGTACTACACCGCGCAACACAGCGCGCGCGCGCGCGCGCGCG 540  
 QY 481 TGGCGGAGGGGCGGTACTACACCGCGCAACACAGCGCGCGCGCGCGCGCGCGCG 540  
 DB 481 TGGCGGAGGGGCGGTACTACACCGCGCAACACAGCGCGCGCGCGCGCGCGCGCG 540  
 QY 541 GCAGCTACACGACAGTGTGTGGGCGCAACCGCGCGAGGTGCGCGCGAGGCCACT 600  
 DB 541 GCAGCTACACGACAGTGTGTGGGCGCAACCGCGCGAGGTGCGCGCGAGGCCACT 600  
 QY 541 GCAGCTACACGACAGTGTGTGGGCGCAACCGCGCGAGGTGCGCGCGAGGCCACT 600  
 DB 541 GCAGCTACACGACAGTGTGTGGGCGCAACCGCGCGAGGTGCGCGCGAGGCCACT 600  
 QY 601 GCGGCGAGGGGCGAGCGCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 DB 601 GCGGCGAGGGGCGAGCGCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 QY 601 GCGGCGAGGGGCGAGCGCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 DB 601 GCGGCGAGGGGCGAGCGCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 QY 661 AGAGCCCTTACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720  
 DB 661 AGAGCCCTTACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720  
 QY 661 AGAGCCCTTACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720  
 DB 661 AGAGCCCTTACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720  
 QY 721 CGGCGGAGGCGAGT 780  
 DB 721 CGGCGGAGGCGAGT 780  
 QY 721 CGGCGGAGGCGAGT 780  
 DB 721 CGGCGGAGGCGAGT 780  
 QY 781 TTTTGT 840  
 DB 781 TTTTGT 840

DB 781 TTTTGT 840  
 QY 841 TCACCTTTTGT 898  
 DB 841 TCACCTTTTGT 898

RESULT 2

ID ABA96418 standard; cDNA; 612 BP.

XX ABA96418;

XX 02-APR-2002 (first entry)

XX Maize PRI-C10 coding sequence SEQ ID NO 3.

XX Maize; pathogen-related; PRI-C10; plant; transgenic; gene; ss.

XX Zea mays.

XX Key location/Qualifiers

XX CDS 1..612

XX FT /\*tag= a

XX FT /product= "PRI-C10"

XX US2001049834-A1.

XX 06-DEC-2001.

XX 10-APR-2001; 2001US-0832320.

XX 10-APR-2000; 2000US-195801P.

XX (CRAN/) CRANE E. H.

XX (CRAN/) CRANE V. C.

XX Crane EH, Crane VC;

XX WPI: 2002-121407/16.

XX P-PSDB: AAM48742.

XX New nucleic acid encoding a pathogen-related protein isolated from  
 PT maize and designated PRI-C10, useful for transforming plants for  
 PT enhanced disease resistance

XX Claim 1; Page 32; 34pp; English.

XX The invention relates to an isolated nucleic acid encoding a  
 CC pathogen-related protein PRI-C10. The nucleic acid is used to transform  
 CC plants for enhanced disease resistance.  
 XX

XX Sequence 612 BP; 80 A; 231 C; 239 G; 62 T; 0 other;

Query Match 68.2%; Score 612; DB 24; Length 612;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-194;  
 Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 ATGGGCGACCTGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 122  
 DB 63 ATGGGCGACCTGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACG 122  
 QY 123 TGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 182  
 DB 123 TGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 182  
 QY 123 TGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 182  
 DB 123 TGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 182  
 QY 183 GCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242  
 DB 183 GCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242  
 QY 183 GCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242  
 DB 183 GCGGTCCTCATGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242  
 QY 243 AGCGGCGACGACGCGAGCGGAGGAGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 302  
 DB 243 AGCGGCGACGACGCGAGCGGAGGAGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 302  
 QY 243 AGCGGCGACGACGCGAGCGGAGGAGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 302  
 DB 243 AGCGGCGACGACGCGAGCGGAGGAGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCG 302

QY	303	GTGGGGGTGGCCCTCCTCGGTGGAGAGCGGAGCTGTGGTTTCGGGCGCCGCGGGAGCGTG	3 62
Db	241	GTGGGGGTGGCCCTCCTCGGTGGAGAGCGGAGCTGTGGTTTCGGGCGCCGCGGGAGCGTG	300
QY	363	GCGCCACACGCGCGCGCAGGGGGGGTGGCGGCTTTCGGGAGAGTGGGGGCGCAGGCCCTACGGC	4 22
Db	301	GCGCCACACGCGCGCGCAGGGGGGGTGGCGGCTTTCGGGAGAGTGGGGGCGCAGGCCCTACGGC	360
QY	423	GCGAAMCAGGGGTGGGCGCAGCTACCGCGCGCCGCCCGCGAGGTGGTGGCTTGGGTG	4 82
Db	361	GCGAAMCAGGGGTGGGCGCAGCTACCGCGCGCCCGCGAGGTGGTGGCTTGGGTG	4 20
QY	483	GCGGAGGGGGGTACTATACACCACGACGACAAACAGTGGCGGGGGGGGGCGGCAATGGCGC	5 42
Db	421	GCGGAGGGGGGTACTATACACCACGACGACAAACAGTGGCGGGGGGGGGCGGCAATGGCGC	4 80
QY	543	ACGTATACGCAAGGTGTGTGGCGCAACCGCGAGGTGGGTGGCGCAGGCCAGGTGC	6 02
Db	481	ACGTATACGCAAGGTGTGTGGCGCAACCGCGAGGTGGGTGGCGCAGGCCAGGTGC	5 40
QY	603	GCGACGGGGCGCACCTACGCTTGTGCTTATACAAACCGCAGGGCAAGGTCAAGGGCAG	6 62
Db	541	GCGACGGGGCGCACCTACGCTTGTGCTTATACAAACCGCAGGGCAAGGTCAAGGGCAG	6 00
QY	663	AGGCCCTACTAG	6 74
Db	601	AGGCCCTACTAG	6 12

RESULT 3  
AAH33782

AC AAH33782;

DE Human colon cancer antigen encoding CDNA SEQ ID NO:838

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

XX

XX  
XX

XX

DR P-PSDB; AAG74351.

PT Nucleic acids enc

PS Claim 1; Page 2787; 9803pp; English.

CC AAH32943 to AAH37195 and AAG75514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders

Sequence 274 BP; 111 A; 48 C; 62 G; 53 T; 0 other;

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
CY

50 Sequence 274 BP; 111 A; 48 C; 62 G; 53 T; 0 other;

```
Query Match      4.88; Score 43; DB 22; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	856	GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	898
Db	211	GGGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	253

RESULT 4  
AAH71507/c  
ID AAH71507 standard; cDNA; 293 BP.

AC AAH71507;

DT 19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 2781.

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN W0200142467-A2

PD 14-JUN-2001.

08-DEC-2000; 2000WO-US33312.

PR 08-DEC-1999; 99US-0169681.

PR 14-MAR-2000; 2000US-0189315.

PR 09-JUN-2000; 2000US-0210600.

[illegible]

XX  
XX  
XX (MILB-) MILLENNIUM PREDICTIO

Pl. Schlegel K, Deeds J, Berge XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer

PS Claim 1; Page 571; 1051pp; English.

The invention relates to novel genes (AAH68172-AAH7383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy.

50 Sequence 293 BP; 54 A; 80 C; 11 G; 148 T; 0 other;



CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 481 BP; 186 A; 94 C; 92 G; 109 T; 0 other;  
Query Match 4.8%; Score 43; DB 23; Length 481;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 856 GCGTAAAAA  
DB 37 GCGTAAAAA  
RESULT 7  
ABV57072/c  
ID ABV57072 standard; cDNA; 550 BP.  
XX  
XX ABV57072;  
XX  
XX 17-SEP-2002 (first entry)  
XX  
XX Human prostate expression marker cDNA 57063.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX  
XX WO200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US05171.  
XX  
XX 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
XX 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
XX 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.  
XX  
XX (MILL-) MILENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
XX Claim 1; Page 10986-10987; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 550 BP; 213 A; 169 C; 34 G; 134 T; 0 other;  
Query Match 4.8%; Score 43; DB 23; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 856 GCGTAAAAA  
DB 145 GCGTAAAAA  
RESULT 8  
ABV56917  
ID ABV56917 standard; cDNA; 551 BP.  
XX  
XX ABV56917;  
XX  
XX 17-SEP-2002 (first entry)  
XX  
XX Human prostate expression marker cDNA 56908.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX  
XX WO200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US05171.  
XX  
XX 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
XX 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
XX 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.  
XX  
XX (MILL-) MILENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
XX Claim 1; Page 10963-10964; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 551 BP; 238 A; 88 C; 74 G; 151 T; 0 other;  
Query Match 4.8%; Score 43; DB 23; Length 551;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 856 GCGTAAAAA 898  
DB 110 GCGTAAAAA 152

RESULT 9  
AAS27462  
ID AAS27462 standard; cDNA; 1077 BP.

AC AAS27462;  
XX  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE cDNA encoding novel signal transduction pathway protein, Seq ID 497.  
XX  
KW Neuroprotective; cytosolic; dermatological; immunosuppressive; tumour;  
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
KW Immune system disorder; Rheumatoid arthritis; Inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS; se;  
KW acquired immune deficiency syndrome.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO20015473-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01312.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0186874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218280.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225265.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236882.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239936.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246539.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.



XX DE Human ORFX ORF2990 polynucleotide sequence SEQ ID NO:5979.  
XX  
KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;  
KW vulnary; antiparkinsonian; nontropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antihematologic;  
KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
KW thrombosis; contraceptive; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200058473-A2.  
XX  
PD 05-OCT-2000.  
XX  
PF 31-MAR-2000; 2000WO-US08621.  
XX  
PR 31-MAR-1999; 99US-0127607.  
XX  
PR 02-APR-1999; 99US-0127636.  
XX  
PR 05-APR-1999; 99US-0127728.  
XX  
PR 30-MAR-2000; 2000US-0540763.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI: 2000-602362/57.  
XX  
DR P-PSDB; AAB43226.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
XX  
PT useful for treating e.g. cancers, proliferative disorders,  
XX  
PT neurodegenerative disorders and cardiovascular disease -  
XX  
XX  
XX Claim 5; Page 5156-5157; 5507pp; English.  
XX  
XX  
XX AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
XX  
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX  
XX sequences have activities such as: cytosolic; hepatotropic; vulnary;  
XX antiparkinsonian; nontropic; neuroprotective;  
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
XX antidiabetic; hypotensive; dermatological; immunosuppressive;  
XX antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
XX antihypertensive; antianaemic. The sequences can be used for determining  
XX the presence of or predisposition to, or preventing or treating  
XX pathological conditions associated with an ORFX-associated disorder. The  
XX nucleic acids can be used to express ORFX proteins in gene therapy  
XX vectors. The proteins and nucleic acids may be used to treat cancers,  
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,  
XX graft vs host disease, cardiovascular disease, diabetes mellitus,  
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance  
XX coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
XX Sequence 1095 BP; 259 A; 333 C; 291 G; 210 T; 2 other;

Query Match 4.8%; Score 43; DB 21; Length 1095;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 856 GGGTAAA 898

DB 1034 GGGTAAA 1076  
|||||  
RESULT 12  
ABL01589  
ID ABL01589 standard; cDNA; 1240 BP.  
XX  
XX ABL01589;  
AC  
XX 15-MAR-2002 (first entry)  
DE  
XX Human secreted protein encoding polynucleotide SEQ ID NO 34.  
XX  
XX Human; nontropic; neuroprotective; cytosolic; dermatological; virocidic;  
KW immunosuppressive; antinflammatory; anti-HIV; antibacterial; vulnary;  
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.  
XX  
XX  
XX Homo sapiens.  
XX  
XX  
XX WO200183510-A1.  
XX  
XX  
XX 08-NOV-2001.  
XX  
XX 26-APR-2001; 2001WO-US13318.  
XX  
XX  
XX 02-MAY-2000; 2000US-201194P.  
XX  
XX 16-JUN-2000; 2000US-212142P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
XX Komatsoulis G, Ruben SM, Rosen CA;  
XX  
XX WPI: 2002-121866/16.  
XX  
XX DR P-PSDB; ABB57417.  
XX  
XX  
XX An isolated nucleic acid molecule encoding a human secreted protein  
XX (SP) is useful in preventing, treating or ameliorating a disorder e.g.,  
XX Alzheimer's disease and cancers -  
XX  
XX  
XX Claim 1; Page 449; 496pp; English.  
XX  
XX  
XX The invention relates to novel genes (ABL01566-ABL01594) and proteins  
XX (ABB57394-ABB57456) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haematolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX  
XX Sequence 1240 BP; 303 A; 303 C; 403 G; 231 T; 0 other;

Query Match 4.8%; Score 43; DB 24; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 856 GGGTAAA 898  
|||||  
DB 1126 GGGTAAA 1168





KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W09839448-A2.  
 XX  
 PD 11-SEP-1998.  
 XX  
 PF 06-MAR-1998; 98WO-0504493.  
 XX  
 PR 02-OCT-1997; 97US-0061060.  
 PR 07-MAR-1997; 97US-0038621.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040162.  
 PR 07-MAR-1997; 97US-0040163.  
 PR 07-MAR-1997; 97US-0040333.  
 PR 07-MAR-1997; 97US-0040334.  
 PR 07-MAR-1997; 97US-0040626.  
 PR 07-MAR-1997; 97US-0043311.  
 PR 11-APR-1997; 97US-0043312.  
 PR 11-APR-1997; 97US-0043313.  
 PR 11-APR-1997; 97US-0043314.  
 PR 11-APR-1997; 97US-0043358.  
 PR 11-APR-1997; 97US-0043569.  
 PR 11-APR-1997; 97US-0043576.  
 PR 11-APR-1997; 97US-0043580.  
 PR 11-APR-1997; 97US-0043669.  
 PR 11-APR-1997; 97US-0043670.  
 PR 11-APR-1997; 97US-0043671.  
 PR 11-APR-1997; 97US-0043672.  
 PR 11-APR-1997; 97US-0043674.  
 PR 23-MAY-1997; 97US-0047590.  
 PR 23-MAY-1997; 97US-0047500.  
 PR 23-MAY-1997; 97US-0047501.  
 PR 23-MAY-1997; 97US-0047502.  
 PR 23-MAY-1997; 97US-0047503.  
 PR 23-MAY-1997; 97US-0047581.  
 PR 23-MAY-1997; 97US-0047582.  
 PR 23-MAY-1997; 97US-0047583.  
 PR 23-MAY-1997; 97US-0047584.  
 PR 23-MAY-1997; 97US-0047585.  
 PR 23-MAY-1997; 97US-0047586.  
 PR 23-MAY-1997; 97US-0047587.  
 PR 23-MAY-1997; 97US-0047588.  
 PR 23-MAY-1997; 97US-0047589.  
 PR 23-MAY-1997; 97US-0047590.  
 PR 23-MAY-1997; 97US-0047592.  
 PR 23-MAY-1997; 97US-0047593.  
 PR 23-MAY-1997; 97US-0047594.  
 PR 23-MAY-1997; 97US-0047595.  
 PR 23-MAY-1997; 97US-0047596.  
 PR 23-MAY-1997; 97US-0047597.  
 PR 23-MAY-1997; 97US-0047598.  
 PR 23-MAY-1997; 97US-0047599.  
 PR 23-MAY-1997; 97US-0047600.  
 PR 23-MAY-1997; 97US-0047601.  
 PR 23-MAY-1997; 97US-0047612.  
 PR 23-MAY-1997; 97US-0047613.  
 PR 23-MAY-1997; 97US-0047614.  
 PR 23-MAY-1997; 97US-0047615.  
 PR 23-MAY-1997; 97US-0047616.  
 PR 23-MAY-1997; 97US-0047618.  
 PR 23-MAY-1997; 97US-0047632.  
 PR 23-MAY-1997; 97US-0047633.  
 PR 06-JUN-1997; 97US-0048964.  
 PR 06-JUN-1997; 97US-0048974.  
 PR 13-JUN-1997; 97US-0049610.

PR 08-JUL-1997; 97US-0051926.  
 PR 16-JUL-1997; 97US-0052874.  
 PR 18-AUG-1997; 97US-0053724.  
 PR 22-AUG-1997; 97US-0056630.  
 PR 22-AUG-1997; 97US-0056631.  
 PR 22-AUG-1997; 97US-0056632.  
 PR 22-AUG-1997; 97US-0056636.  
 PR 22-AUG-1997; 97US-0056637.  
 PR 22-AUG-1997; 97US-0056662.  
 PR 22-AUG-1997; 97US-0056664.  
 PR 22-AUG-1997; 97US-0056845.  
 PR 22-AUG-1997; 97US-0056862.  
 PR 22-AUG-1997; 97US-0056864.  
 PR 22-AUG-1997; 97US-0056872.  
 PR 22-AUG-1997; 97US-0056874.  
 PR 22-AUG-1997; 97US-0056875.  
 PR 22-AUG-1997; 97US-0056876.  
 PR 22-AUG-1997; 97US-0056877.  
 PR 22-AUG-1997; 97US-0056878.  
 PR 22-AUG-1997; 97US-0056879.  
 PR 22-AUG-1997; 97US-0056880.  
 PR 22-AUG-1997; 97US-0056881.  
 PR 22-AUG-1997; 97US-0056882.  
 PR 22-AUG-1997; 97US-0056884.  
 PR 22-AUG-1997; 97US-0056886.  
 PR 22-AUG-1997; 97US-0056887.  
 PR 22-AUG-1997; 97US-0056888.  
 PR 22-AUG-1997; 97US-0056889.  
 PR 22-AUG-1997; 97US-0056892.  
 PR 22-AUG-1997; 97US-0056893.  
 PR 22-AUG-1997; 97US-0056894.  
 PR 22-AUG-1997; 97US-0056903.  
 PR 22-AUG-1997; 97US-0056908.  
 PR 22-AUG-1997; 97US-0056910.  
 PR 22-AUG-1997; 97US-0056911.  
 PR 22-AUG-1997; 97US-0056912.  
 PR 05-SEP-1997; 97US-0057650.  
 PR 05-SEP-1997; 97US-0057669.  
 PR 05-SEP-1997; 97US-0057761.  
 PR 12-SEP-1997; 97US-0058785.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Bedharik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;  
 PI Feng P, Ferrie AM, Fischer CL, Florence RA, Greene JM, Hu JS;  
 PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;  
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
 XX  
 DR WPI: 1998-506364/43.  
 DR P-PSDB: AAW74810.  
 XX  
 PT New isolated human genes and the secreted polypeptide(s) they encode  
 PT - useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 PS  
 XX  
 Claim 1; Page 311-312; 721pp; English.  
 CC This sequence represents a nucleic acid molecule designated Gene 80 from  
 CC the human CDNA clone HWMR27 (deposited as clone ATCC 97900 and ATCC  
 CC 20904b) which encodes a secreted human protein. The gene can be used to  
 CC generate fusion proteins by linking to the gene to a human  
 CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of  
 CC the fused protein as compared to the human protein only.  
 CC The invention relates to 186 novel genes and their firmaments (nucleic  
 CC acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the 186  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAV59511 for described uses).

Sequence 1821 BP; 545 A; 429 C; 343 G; 497 T; 7 other;

Query Match 4.8%; Score 43; DB 19; Length 1821;

Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 GCGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 898

DB 1661 GCGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1703

Search completed: February 22, 2003, 04:15:14  
Job time : 248.123 secs

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100.

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 03:42:31 ; Search time 1698.6 Seconds

(without alignments)  
10485.621 Million cell updates/sec

Title: US-09-832-320-3

Perfect score: 612

Sequence: 1 atggcgcaactcgcgagcca.....aggcgagagccctactag 612

Scoring table: OLIGO\_NTG  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 25

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vl:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.inv:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.pln:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Match | Query length | DB ID | Description |
|------------|-------|-------|--------------|-------|-------------|
| 1          | 40    | 6.5   | 144973       | 2     | AC096689    |
| 2          | 40    | 6.5   | 150928       | 2     | AC103550    |

#### ALIGNMENTS

RESULT 1  
AC096689 144973 bp DNA linear HTG 09-JAN-2002  
LOCUS DEFINITION Oryza sativa chromosome 3 clone OSJNBa0027J18, \*\*\* SEQUENCING IN  
AC096689 PROGRESS \*\*\*, 6 ordered pieces.  
AC096689 AC096689.3 GI:18092993  
VERSION HTG, HTGS\_PHASE2.  
KEYWORDS  
SOURCE ORGANISM  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartioideae; Oryzae; Oryza.  
1 (bases 1 to 144973)  
Buell, R., Hsiao, J., Zismann, V., Moffat, K.M., Hill, J.,  
Gansberger, K., Burgess, S., Jarral, B., Shvartsbeyn, M., Brenner, M.,  
Ciecko, A., Pal, G., Vanaken, S., Hansen, C., Ulterbach, T.,  
Feldblum, T., Khalak, H.G., Yuan, Q., Quackenbush, J., White, O.,  
Salzberg, S. and Fraser, C.  
Oryza sativa ssp. japonica cv. Nipponbare OSJNBa0027J18 BAC genomic  
sequence  
Unpublished  
2 (bases 1 to 144973)  
Buell, R.  
Direct Submission  
Submitted (22-SEP-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
On Jan 9, 2002 this sequence version replaced gi:17530731.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 29894: contig of 29894 bp in length  
29895 29894: gap of unknown length  
29995 70108: contig of 40114 bp in length  
70109 70208: gap of unknown length  
70209 79598: contig of 9390 bp in length  
79599 79698: gap of unknown length  
79699 106423: contig of 26723 bp in length  
106424 106523: gap of unknown length  
106524 139626: contig of 33103 bp in length  
139627 144973: gap of unknown length  
139727 144973: contig of 5247 bp in length.

#### FEATURES

##### source

1. 144973  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="3"  
/clone="OSJNBa0027J18"

BASE COUNT 38927 a 32302 c 32466 g 40767 t 511 others  
ORIGIN

Pred. No. is the number of results predicted by chance to have a

Query Match 6.5%; Score 40; DB 2; Length 144973;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 TACAACCCGCGACGCAAGTGCAGGCGCCAGCCCTACT 610  
 |||||  
 Db 136469 TACAACCCGCGACGCAAGTGCAGGCGCCAGCCCTACT 136508

RESULT 2  
 AC103550/c  
 LOCUS  
 DEFINITION Oryza sativa chromosome 3 clone OSJNBa0079G12, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 17 unordered pieces.  
 AC103550  
 AC103550.6 GI:21322019  
 HTG: HTGS\_PHASE1.  
 SOURCE  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoae; Oryza.  
 1 (bases 1 to 150928)  
 Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Tallon,L., Gausberger,K.,  
 Kim,M., Overton II,L., Bera,J., Tsirlin,T., Krol,M., Jarrohl,B.,  
 Jin,S., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S.,  
 Uterbach,T., Feldblyum,T., Yang,Q., Haas,B., Sub,B., Peterson,J.,  
 Quackenbush,J., White,O., Salzberg,S. and Fraser,C.  
 Oryza sativa ssp. japonica cv. Nipponbare OSJNBa0079G12 BAC genomic  
 sequence  
 Unpublished  
 2 (bases 1 to 150928)  
 Buell,R.  
 Direct Submission  
 Submitted (27-NOV-2001) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 3 (bases 1 to 150928)  
 Buell,R.  
 Direct Submission  
 Submitted (04-JUN-2002) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 On Jun 4, 2002 this sequence version replaced gi:20153306.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 17 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 5616: contig of 5616 bp in length  
 \* 5617 5663: gap of unknown length  
 \* 5664 31941: contig of 26278 bp in length  
 \* 31942 31988: gap of unknown length  
 \* 31989 42994: contig of 11006 bp in length  
 \* 42995 43041: gap of unknown length  
 \* 43042 45430: contig of 2389 bp in length  
 \* 45431 45477: gap of unknown length  
 \* 45478 62104: contig of 16627 bp in length  
 \* 62105 62151: gap of unknown length  
 \* 62152 64960: contig of 2809 bp in length  
 \* 64961 80507: gap of unknown length  
 \* 80508 80503: contig of 15496 bp in length  
 \* 80504 80550: gap of unknown length  
 \* 80551 85002: contig of 4452 bp in length  
 \* 85003 85048: gap of unknown length  
 \* 85049 95622: contig of 10573 bp in length  
 \* 95623 95669: gap of unknown length  
 \* 95670 109958: contig of 14289 bp in length  
 \* 109959 110005: gap of unknown length  
 \* 110006 124211: contig of 14206 bp in length  
 \* 124212 124258: gap of unknown length

124259 131728: contig of 7470 bp in length  
 \* 131729 131775: gap of unknown length  
 \* 131776 133914: contig of 2139 bp in length  
 \* 133915 133962: gap of unknown length  
 \* 133962 141548: contig of 7887 bp in length  
 \* 141549 141595: gap of unknown length  
 \* 141596 144303: contig of 2708 bp in length  
 \* 144304 144349: gap of unknown length  
 \* 144350 148834: contig of 4485 bp in length  
 \* 148835 148880: gap of unknown length  
 \* 148881 150928: contig of 2048 bp in length.

FEATURES  
 source  
 1. 150928  
 /organism="Oryza sativa"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /chromosome="3"  
 /clone="OSJNBa0079G12"  
 /note="japonica cultivar-group"  
 BASE COUNT 42065 a 34725 c 33029 g 40330 t 779 others  
 ORIGIN

Query Match 6.5%; Score 40; DB 2; Length 150928;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 TACAACCCGCGACGCAAGTGCAGGCGCCAGCCCTACT 610  
 |||||  
 Db 68225 TACAACCCGCGACGCAAGTGCAGGCGCCAGCCCTACT 68186

Search completed: February 22, 2003, 05:29:30  
 Job time : 1908.6 secs